BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/12, 15/86, C12P 21/02
C07K 13/00, C12Q 1/68, 1/42

A2

(11) International Publication Number: WO 90/00607

(43) International Publication Date: 25 January 1990 (25.01.90)

(21) International Application Number: PCT/US89/02761

(22) International Filing Date: 23 June 1989 (23.06.89)

(30) Priority data:

216,888	8 July 1988 (08.07.88)	US
230,761	10 August 1988 (10.08.88)	US
260,807	21 October 1988 (21.10.88)	US

(71) Applicant: CETUS CORPORATION [US/US]; 1400 Fifty-Third Street, Emeryville, CA 94608 (US).

(72) Inventors: WONG, Gail, L. ; 11255 Elvessa Street, Oakland, CA 94605 (US). MCCORMICK, Francis, P. ; 921 Ramona Avenue, Albany, CA 94706 (US). MARTIN, George ; 2346 Woolsey Street, Berkeley, CA 94705 (US). RUBINFELD, Bonnee ; 3403 Claridge Drive, Danville, CA 94526 (US). O'ROURKE, Edward, C. ; 7321 Skyline Boulevard, Oakland, CA 94611 (US). CLARK, Robin ; 3736 Woodruff Avenue, Oakland, CA 94602 (US). KOTHS, Kirston, E. ; 2646 Mira Vista Drive, El Cerrito, CA 94530 (US). HALENBECK, Robert, F. ; 136 Spring Grove Avenue, San Rafael, CA 94901 (US). TRAHEY, Mary, M. ; 5933 Chabot Road, Oakland, CA 94618 (US).

(74) Agent: HALLUIN, Albert, P.; Cetus Corporation, 1400 Fifty-Third Street, Emeryville, CA 94608 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), BG, CH (European patent), DE (European patent), DK, FI, FR (European patent), GB (European patent), HU, IT (European patent), JP, LU (European patent), NL (European patent), NO, SE (European patent).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: GAP GENE SEQUENCES AND DIAGNOSTIC USES THEREOF

ILE MET PRO GLU GLU GLU TYR SER GLU PHE LYS

```
ATC ATG CCC GAG CAG GAG TAC TCC GAG TTC AAG
  T     A   AGA   A   T   T   A   T   A
5 A     T                 A
        G                 G
                          AGC
                          T
```

(57) Abstract

Guanosine triphosphatase activating protein (GAP) DNA sequences are described that are useful as cancer diagnostics, particularly to detect cancer cells that express the ras oncogene protein p21 by measuring the level of GAP gene expression or amplification.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MR	Mauritania
BE	Belgium	GA	Gabon	MW	Malawi
BF	Burkina Faso	GB	United Kingdom	NL	Netherlands
BG	Bulgaria	HU	Hungary	NO	Norway
BJ	Benin	IT	Italy	RO	Romania
BR	Brazil	JP	Japan	SD	Sudan
CA	Canada	KP	Democratic People's Republic of Korea	SE	Sweden
CF	Central African Republic			SN	Senegal
CG	Congo	KR	Republic of Korea	SU	Soviet Union
CH	Switzerland	LI	Liechtenstein	TD	Chad
CM	Cameroon	LK	Sri Lanka	TG	Togo
DE	Germany, Federal Republic of	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

GAP GENE SEQUENCES AND DIAGNOSTIC USES THEREOF

Field of the Invention

This invention relates generally to the field of oncology, and particularly to compositions useful in diagnostic testing for cancer. More specifically, the invention concerns DNA sequences, and compositions containing the same, that can be employed as cancer diagnostics.

Background of the Invention

Several genes have been identified that are thought to play a role in regulating normal cell growth. A subset of these genes, termed ras, consists of at least three members, N-ras, H-ras, and K-ras2. Altered forms of ras, termed oncogenes, have been implicated as causative agents in cancer. Both the normal cellular genes and the oncogenes encode chemically related proteins, generically referred to as p21.

Ras oncogenes, and their normal cellular counterparts, have been cloned and sequenced from a variety of species. Comparison of the structure of these two genes has revealed that they differ by point mutations that alter the amino acid sequence of the p21 protein. Naturally occurring mutations in the ras oncogenes have been identified in codons 12, 13, 59, and 61. In vitro mutagenesis work has shown that mutations in codon 63, 116, 117 and 119 also result in transforming activity. The most frequently observed mutation which converts a normal cellular ras gene into its oncogenic counterpart is a substitution of glycine at position 12 by any other amino acid residue, with the exception of proline. Transforming activity is also observed if glycine is deleted, or if amino acids are inserted between alanine at position 11 and glycine at position 12.

Mutations at position 61 also play an important role in the generation of ras oncogenes. Substitution of glutamine for any other amino acid, except proline or glutamic acid in the cellular ras gene yields ras oncogenes with transforming activity.

In relation to normal cellular ras genes and their oncogenic counterparts, there are at least four known retroviral ras oncogenes which exhibit transforming activity. Unlike their non-retroviral analogues, the retroviral genes exhibit two mutations. The biological significance of these double mutations is at present

unclear.

Both the normal ras and oncogenic p21 proteins, regardless of their phylogenetic origin, bind guanine nucleotides, GTP and GDP, and possess intrinsic GTPase activity. See Temeles et al., 1985 Nature, 313:700. The significance of these biochemical properties to the biological activities of the ras proteins has been demonstrated as follows: first, microinjection of anti-ras antibodies that interfere with guanine nucleotide binding reverses the malignant phenotype of NIH 3T3 cells transformed by ras oncogenes. See Clark et al., 1985 Proc. Natl. Acad. Sci. U.S.A., 82:5280 and Feramisco et al., 1985 Nature, 314:639. Second, ras oncogenic proteins that exhibit mutations which result in the inability of p21 to bind guanine nucleotides do not transform NIH 3T3 cells. Willumsen et al., 1986 Mol. Cell. Biol., 6:2646. Third, some ras oncogenes produce p21 proteins that have much reduced GTPase activity compared to their normal cellular counterparts. The biological role of GTPase activity associated with either ras or its oncogenic counterpart remains unknown.

Recently a cytoplasmic factor has been identified which stimulates normal ras p21 GTPase activity, but does not effect GTPase activity associated with the oncogenic mutants. See M. Trahey and F. McCormick, 1987 Science, 238:542. The activity has been associated with a protein, termed GAP, which is the acronym for GTPase activating protein. GAP is thought to be a cytoplasmic protein but is presumably capable of moving from the cytosol to the plasma membrane where it interacts with p21.

As alluded to above, ras oncogenes have been implicated in the development of a variety of tumors, and have been shown to be involved in about 10-40% of the most common forms of human cancer. See H. Varmus, 1984 Annual Rev. Genetics, 18:553 and M. Barbacid, 1986, in Important Advances in Oncology, ed. B. DeVita, S. Helman, S. Rosenberg, pages 3-22, Philadelphia:Lippincott. For example, ras oncogenes have been consistently identified in carcinomas of the bladder, colon, kidney, liver, lung, ovary, pancreas and stomach. They also have been identified in hematopoietic tumors of lymphoid and myeloid lineage, as well as in tumors of mesenchymal origin. Furthermore, melanomas, teratocarcinomas, neuroblastomas, and gliomas have also been shown to possess ras oncogenes.

Considering the possible association of ras oncogenes and cancer, there has

been considerable work focused on diagnostic tests for detecting the presence of the oncogene product, p21, or the mutant oncogenes. Early tests, which are still employed in many instances, identify the presence of ras oncogenes in transfection assays which identify p21 by its ability to transform NIH 3T3 cells. See Lane et al., 1981 Proc. Natl. Acad. Sci. USA, 78:5185 and B. Shilo, and R. A. Weinberg, 1981 Nature, 289:607. This method is insensitive, laborious, and to be performed adequately, requires a skilled laboratory technician.

A second diagnostic method centers around oligonucleotide probes to identify single, point mutations in genomic DNA. This technique is based on the observation that hybrids between oligonucleotides form a perfect match with genomic sequences, that is, non-mutated genomic sequences are more stable than those that contain a single mismatch. An example of the latter is a point mutation in p21 associated with the ras oncogenes. Although this technique is clearly more sensitive and easier to perform than the transfection assay, it is nevertheless also cumbersome to perform. This is because there are theoretically almost 100 base substitutions which can yield ras oncogenes. Thus, in order to be able to detect these substitutions, multiple oligonucleotide probes must be employed containing each of the three possible substitutions at a particular residue. See Bos et al., 1985 Nature, 315:726 and Valenzuela et al., 1986 Nuc. Acid Res., 14:843.

In addition to the transfection and oligonucleotide assays, additional nucleic acid hybridization techniques have been developed to identify ras oncogenes. One such method is based on the unusual electrophoretic migration of DNA heteroduplexes containing single based mismatches in denaturing gradient gels. See Myers et al., 1985 Nature, 313:495. This technique only detects between about 25-40% of all possible base substitutions, and requires a skilled technician to prepare the denaturing gradient gels. More sensitive techniques which are refinements of this technique are described by Winter et al., 1985 Proc. Natl. Acad. Sci. USA, 82:7575 and Myers et al., 1985 Science, 230:1242.

Immunologic approaches have been taken to detect the product of the ras oncogenes. Polyclonal or monoclonal antibodies have been generated against the intact ras oncogene p21, or against chemically synthesized peptides having sequences similar to oncogene p21, or the non-transforming counterpart. See U.S. Patent Application Serial No. 938,581; EP Patent Publication 108,564 to Cline et al.; Tamura et al., 1983 Cell, 34:587; PCT Application WO/84/01389 to Weinberg et al. For the most part antibodies have been disappointing as diagnostic tools with which to identify ras oncogenic p21 in human tissue sections. This is because either the antibodies that have been generated to date recognize the normal cellular ras protein as well as the oncogenic protein, or, in those instances in which a monoclonal antibody has been generated that specifically recognizes the oncogenic protein, non-specific staining of tumor biopsies is still observed.

While ras oncogenic p21 is an effective tumorigenic agent, recent studies have shown that normal ras p21 can induce the malignant phenotype. See Chang et al., 1982 Nature, 297:7479 and Pulciani et al., 1985 Mol. Cell. Biol., 5:2836. For example, transfection of normal H-ras DNA has been shown to induce malignant transformation. It is further noteworthy that normal ras gene amplification has been observed in several human tumors, and has an apparent incidence of about 1%. Pulciani, et al., above; Yokota et al., 1986 Science, 231:261. The various diagnostic tests used to detect ras oncogenes or oncogenic p21 have been applied to the detection of normal ras p21 with similar limited success.

It should be apparent from the foregoing that while there are a number of diagnostic methods for determining the presence of ras oncogenes, or their transforming proteins, there is still a need for fast and reliable diagnostic methods that will permit the routine identification of a wide variety of ras related tumors.

Summary of the Invention

In accordance with the present invention, DNA sequences are described that are useful as diagnostics for cancers arising from the expression of normal cellular or oncogenic ras genes.

A second aspect of the invention is a description of GAP cDNA sequences, and methods for isolating and identifying the same, that are useful in cancer diagnosis.

A third aspect of the invention is a description of a full length GAP cDNA sequence that encodes a molecule with a molecular weight of about 116,000, and cDNAs that encode shorter GAP molecules resulting from differential splicing.

A fourth aspect of the invention is a description of the procedures employed to realize the expression of GAP cDNA in bacteria or insect cells.

A fifth aspect of the invention consists of diagnostic methods for detecting cancer using GAP gene sequences.

Further aspects of the invention will become apparent upon consideration of the following description of the invention.

Brief Description of the Drawings

Figure 1 shows the TSK phenyl column elution profile and silver staining of SDS PAGE fractions thereof.

Figure 2 shows the SDS gel profile of GAP purified by a three-step chromatographic scheme consisting of cation, and anion chromatography, followed by a second cation chromatographic step.

Figure 3 shows the GAP amino acid sequence used to generate DNA probes that were used to identify the lambda gt11 clone, GAP 6. Also shown is the corresponding DNA encoding sequence with possible codon redundancies.

Figure 4 shows the DNA probes used to identify GAP 6.

Figure 5 presents the DNA and amino sequence of lambda clone, clone 101.

Figure 6 presents the DNA sequence of the lambda clones, clone 16 and clone "Sleepy".

Figure 7 shows the results of GAP assays conducted in the presence of lysates prepared from lambda lysogens of clones 7 and 101 (top panel); and of Sf9 cell lysates transfected with pAcC12-GAP 5 (bottom panel).

Figure 8 shows the construction of pAcC12.

Detailed Description of the Invention

A better understanding of the invention described herein will be realized by

providing a brief description of some of the materials and methods used in the invention.

The normal cellular ras gene and its oncogenic counterparts are defined as described by N. Barbacid, 1987 Ann. Rev. Biochem., 56:779. Similarly, the proteins encoded by these genes are also defined as described by Barbacid. Moreover, it will be appreciated that fragments of normal cellular p21 that bind GTP, and exhibit GAP stimulated GTPase activity are intended to come within the definition of ras p21.

GAP is the acronym for guanine triphosphatase activating protein, and is defined as a protein having a molecular weight and amino acid sequence as described herein, and that has the further properties of stimulating GTPase activity of normal cellular ras p21, while having little or no stimulatory activity when combined with oncogenic ras p21 proteins and GTP. Of course, it will be understood by those skilled in the art that GAP may also exist as aggregates or multimers under certain conditions, and these forms are intended to come within the scope of the definition. Moreover, the definition is further intended to cover fragments of GAP that exhibit activity. Exemplary of such a fragment is a molecule having a reduced subunit molecular weight of about 35,000 as shown herein.

It will further be appreciated with regard to the chemical structure of GAP, that its precise structure may depend on a number of factors. As all proteins contain ionizable amino and carboxyl groups, it is, of course, apparent that GAP may be obtained in acid or basic salt form, or in neutral form. It is further apparent, that the primary amino acid sequence may be augmented by derivatization using sugar molecules (glycosylation) or by other chemical derivatizations involving covalent or ionic attachment to GAP with, for example, lipids, phosphate, acetyl groups and the like, often occurring through association with saccharides. These modifications may occur in vitro or in vivo, the latter being performed by a host cell through post-translational processing systems. It will be understood that such modifications, regardless of how they occur, are intended to come within the definition of GAP so long as the activity of the protein, as defined herein, is not significantly altered.

As used herein, "chromatography" is defined to include application of a solution containing a mixture of compounds to an adsorbent, or other support material which is eluted, usually with a gradient or other sequential eluant. Material eluted from the support matrix is designated eluate. The sequential elution is most routinely performed by isolating the support matrix in a column and passing the eluting solution(s), which changes affinity for the support matrix, either stepwise or preferably by a gradient, through the matrix. It will be appreciated that encompassed within the definition "chromatography" is the positioning of the support matrix in a filter and the sequential administering of eluant through the filter, or in a batch-mode.

The phrase "hydrophobic interaction matrix" is defined to mean an adsorbent that is a hydrophobic solid such as polystyrene resin beads, rubber, silica-coated silica gel, or crosslinked agarose sufficiently substituted with hydrophobic functional groups to render the material hydrophobic. Alkyl substituted agarose and aryl substituted agarose such as, for example, phenyl or octyl agarose are representative hydrophobic materials. Mixtures of materials that are chromatographically separated on a hydrophobic interaction chromatography matrix are generally first adsorbed to the matrix in a high salt solution, and subsequently desorbed from the matrix by elution in a low salt solution, or a hydrophobic solvent such as a polyol.

"Anion exchange matrix" is defined to mean a solid or gel support matrix that is charged in aqueous solutions. The support matrix may be agarose sufficiently substituted with amine functional groups to have a net charge in aqueous solutions. The material to be adsorbed is generally bound to the anion exchange matrix in a low salt solution and is generally eluted from the anion exchange matrix in a high salt eluant containing anions such as chloride ion which bind to the anion exchange matrix and displace the adsorbed material.

By the phrase "high salt concentration conditions" is meant an aqueous solution wherein an ionic substance is present to create conditions of high ionic strength. Ionic strength is defined as is generally understood in the art and can be calculated from the putative concentrations of the various ions placed in solution modified by their activity coefficient. High salt concentrations that are routinely employed are typified by solutions containing high concentrations of ammonium sulfate; however, other salts, such as sodium chloride, potassium chloride, sodium sulfate, sodium nitrate, or sodium phosphate may also be employed.

The definition of "affinity chromatography" is understood to be similar to that of Wilchek et al., 1984 Methods in Enzymology, 104:3. In its broadest intended definition, "affinity chromatography" is a "method of purification based on biological recognition". Briefly, the procedure involves coupling a ligand to a solid support, and contacting the ligand with a solution containing therein a ligand recognition molecule which binds to the ligand. Subsequently, the ligand recognition molecule is released from the ligand and isolated in pure form. It will be understood that a variety of ligands can be employed in affinity chromatography as discussed by Wilchek, et al., and examples of these include lectins, antibodies, receptor-binding proteins and amino acids.

"Cells" or "recombinant host" or "host cells" are often used interchangeably as will be clear from the context. These terms include the immediate subject cell, and, of course, the progeny thereof. It is understood that not all progeny are exactly identical to the parental cell, due to chance mutations or differences in environment. However, such altered progeny are included when the above terms are used.

General Description

The instant invention provides a description of DNA sequences that encode GAP, and materials and methods for identifying and isolating the same. The DNA sequences, or fragments derived therefrom, are useful as cancer diagnostics, being particularly useful to diagnose for ras p21 related cancers. The identification and isolation of the instant GAP DNA sequences is facilitated by the availability of DNA oligonucleotide probes substantially homologous to the GAP sequence. Because such probes were generated based on a knowledge of the partial amino acid sequence of GAP, the order of discussion of the invention will be: purification of GAP; methods of assaying GAP; the partial amino acid sequence of GAP; cloning of GAP using GAP probes based on the amino acid sequence and the identification of GAP DNA sequences in a cDNA library, along with subcloning of the sequences. In this section is also described the expression of the GAP sequences along with methods of using the same to diagnose for cancer.

GAP Purification

Guanosine triphosphatase activating protein, or GAP, is widely expressed in higher eukaryotes. GAP has been detected in cell extracts from human and mouse normal tissues including brain, liver, placenta, B cells, and platelets. It has additionally been found in non-transformed cell cultures including NIH 3T3, as well as transformed cell lines, including human mammary cancer cells (MCF-7), retinoblastoma cells (Y79), and Wilm's tumor (G401). GAP is also present in insect cells such as, for example, Spodoptera fragipedra. From many of these cells or tissues, GAP may be isolated, albeit with minor variations in the purification protocols and the like.

The general scheme for GAP isolation and purification consists of releasing the molecule from the cytoplasm of appropriate cells, tissues or organs, followed by removing insoluble material and subjecting the soluble GAP fraction to cation exchange chromatography, followed by a second chromatographic step wherein the

eluant from the cation exchanger is passed over an anion exchanger. GAP is eluted from the anion exchanger, and further purified by subjecting it to a third chromatographic step, either hydrophobic chromatography, or a second cation exchange step.

More specifically, GAP is prepared by releasing the molecule from the cytosol using any number of techniques including freeze thawing, sonication, mild detergent extraction, etc. This procedure is preferably carried out in a physiologically buffered solution containing one or more protease inhibitors. Moreover, to further inhibit protease activity, especially those proteases that rely on metal ions for activity, the extraction solution may contain metal ion chelators. The preferred extraction solution is a physiologically balanced salt solution containing the chelators ethyleneglycoltrichloroacetic acid (EGTA), or ethylenediaminetrichloroacetic acid (EDTA), plus the protease inhibitor phenylmethylsulfonylfluoride (PMSF). The metal ion chelator(s), as well as the protease inhibitor(s) are present at concentrations that effectively inhibit proteolysis, preferably about 5 mM and 100 μM, respectively. However, it will, of course, be appreciated by those skilled in the art that since the types and amounts of proteases vary depending on the starting material used to extract GAP, the concentrations that the protease inhibitors or chelators are used at, if indeed used at all, will also vary.

The mixture containing GAP is clarified by centrifugation, or in other ways to remove insoluble material from the aqueous cytosol fraction. If the cytosol fraction contains low amounts of GAP it can be concentrated by any one of several techniques well known to those skilled in the art, including high salt precipitation, such as, for example, with ammonium sulfate, or by ultra filtration. If GAP is concentrated by precipitation, it is preferably subsequently resuspended in a suitable physiologically balanced salt solution containing protease inhibitor(s) and preferably about 0.1% of a nonionic detergent, such as NP40. This solution is then prepared for ion exchange chromatography by dialyzing it against a compatibly buffered chromatographic solution, preferably containing millimolar phosphate, a metal ion chelator, a reducing agent, and a protease inhibitor. Additionally, because GAP

activity is stimulated by the presence of divalent cations such as magnesium chloride, it may also be present in the solution. The pH of the solution is preferably about 6.0.

The GAP dialyzate is then subjected to chromatographic purification consisting preferably of three steps. The first involves purification using an ion exchange chromatographic step compatible with the GAP extraction buffer. Since the preferred extraction buffer contains phosphate, the initial step is purification of GAP by cation exchange chromatography. The second consists of ion exchange chromatography wherein the ion exchange matrix has the opposite ion binding capacity from that of the first ion exchanger employed.

Thus, the preferred purification scheme will consist of applying the phosphate solution containing GAP to a cation exchanger, and eluting GAP therefrom, preferably using solutions which alter the pH or conductivity of the solution. More preferably, GAP will be eluted by applying either a gradient or non-gradient salt solution, and most preferably will be eluted using a linear gradient of sodium chloride over the range of about 0-0.6 molar.

The preferred cation exchanger is a SP-cellulose cation exchanger. Such are commercially available from AMF Molecular Separations Division, Meridian, CT under the brand name ZetaPrep SP cartridges. The SP-cellulose cation exchanger is an elastic 3-dimensional network composed of cellulosic backbones cross-linked with vinyl polymer containing pendant sulfopropyl functional groups. The matrix is preferably adapted for radial flow passage of the GAP solution. The flow rate of the solution through the matrix will depend upon the size and geometry of the matrix used. It will be apparent to those skilled in the art, however, that care should be taken to avoid exceeding the unit capacity of the matrix with GAP. If the capacity is exceeded, GAP will not be totally retained and excess unretained GAP will be present in the effluent. The capacity of the matrix to retain GAP can be monitored by assaying for GAP in the effluent using one of the assays described below.

Fractions containing GAP are prepared for the second chromatographic step, that is, anion exchange chromatography. This consists of combining the fractions and adjusting the solution to a pH, and ionic strength compatible with anion exchange chromatography. A variety of anion exchangers are available, and depending on the type employed, the concentrations of these reagents will vary. DEAE-Sepharose or

TSK-DEAE-5-PW may be employed. The general procedures for preparing and using these matrices are known to those skilled in the art. The preferred anion exchanger is TSK-DEAE-5-PW matrix. It is prepared by equilibrating it with a solution containing chloride ions at a pH of 8.5. More preferably, the solution will consist of Tris hydrochloride, pH 8.5 plus a reducing agent, a metal chelator, magnesium chloride, and a protease inhibitor. The concentrations of the metal chelator and protease inhibitor will vary and depend on how extensively GAP is proteolyzed, and whether the proteases responsible are activated by metal ions. The concentration of monovalent cations, such as magnesium chloride and reducing agent can be determined empirically by monitoring GAP activity. Those concentrations which maintain the highest activity will be utilized. Generally, it is preferred that magnesium chloride and the reducing agent be present in the range of about 0.5-1 mM, and 0.1-1 mM, respectively.

The solution is then passed through the anion exchange matrix whereupon GAP binds to the matrix. GAP is subsequently eluted from the matrix using solutions which alter the pH or conductivity. The preferred elution method consists of eluting GAP using a linear salt gradient ranging from 0-0.6 molar sodium chloride. The purity and activity of GAP so obtained can be monitored by the GTPase assay described below, and by sodium dodecyl sulfate polyacrylamide gel electrophoresis run under reducing conditions. Using these techniques it was determined that GAP has a molecular weight of about 115,000-120,000 daltons.

The third chromatographic step consists of applying, after the anion exchange chromatography, either a second cation exchange step, or a hydrophobic interaction chromatographic step. The most preferred purification scheme utilizes a second cation exchange step. Application of either of these methods will generally increase the purity of GAP to about 95%. If a cation exchange column is chosen, the materials and methods described above are similarly applicable here. Generally, this will consist of decreasing the salt concentration present in the anion column eluates and adjusting the pH to about 6.0. Here, as in the initial cation chromatographic step, several different types of cation exchange matrices can be employed; however, the preferred matrix is a SP-TSK column which is run under high pressure. If hydrophobic chromatography is selected, the ionic strength of the eluate from the anion exchanger should be increased to be compatible with hydrophobic interaction

chromatography. The solution can then be passed through a hydrophobic interaction chromatographic matrix, and eluted using techniques known in the art, including decreasing the salt concentration, or eluting with a chaotropic agent. Either of the latter solutions may be used alone, or in combination.

A variety of hydrophobic interaction chromatographic matrixes may be utilized. Generally, the materials and methods for utilizing hydrophobic chromatography are described by S. Shaltie, 1984 Methods in Enzymology, 104:69. While it is apparent there are many hydrophobic chromatographic materials and methods that may be employed to purify GAP, phenyl Sepharose is preferred, and it is further preferred that the chromatography be employed under high pressure. The general procedures for forming high pressure liquid chromatography involving a phenyl derivatized matrix are described by F. Regmaer, 1983 Methods in Enzymology, 91:137. The preferred phenyl derivatized matrix is available commercially from Bio-Rad Corporation, and is sold under the trade name Biogel TSK phenyl-5-PW.

It will be additionally appreciated by those skilled in the art that an alternative purification scheme may consist of a cation and anion chromatographic exchange step, followed by an affinity chromatographic step. This may be achieved by binding GAP to one or more plant lectins having a known carbohydrate specificity compatible with carbohydrates which may be present on GAP, or by binding GAP to anti-GAP antibodies. In either event, GAP can then be released from the affinity matrix using the appropriate sugar if the matrix is composed of a lectin, or by pH or chaotropic agents if the matrix is composed of antibody.

Because GAP is a protease-sensitive molecule that is broken down into lower molecular weight species having GAP activity, in a preferred embodiment of the invention the entire purification procedure is carried out rapidly in the cold to reduce protease activity. In general, this temperature is in a range below 10°C, with a preferred temperature range being about 2-8°C. Most preferred is a temperature of about 4°C.

Finally, it should be noted that while the preferred applications of the ion exchange materials described herein are in a column format, it will be appreciated that they may also be used in batch format as well.

A preferred embodiment purification scheme consists of isolating GAP from human placentas as follows.

GAP was isolated from 300 g of human placentas by the following three-step chromatographic procedure. Placentas were obtained shortly after delivery, and kept on ice until they were processed. After it was determined by standard tests that the placentas were free of HIV antibodies, they were processed as follows. The initial step consisted of mechanically removing connective tissue, and ridding the placentas of excess blood by multiple soakings in phosphate buffered saline (PBS). The placentas were then fragmented by freezing the tissue at -70°C, followed by placing the tissue in solution of PBS containing 5 mM EGTA, 100 μM PMSF and disrupting the tissue in a blender until a uniform, fine suspension was apparent. The suspension was centrifuged at 100,000 x g to remove insoluble debris, the supernatant removed and the proteinaceous material therein precipitated with 40% ammonium sulfate. The ammonium sulfate was removed, and the precipitated proteins resuspended in PBS containing 0.1% NP40 and 100 μM PMSF. This solution was immediately dialyzed against 20 mM potassium phosphate, 1 mM $MgCl_2$, 5 mM EGTA, 0.1 mM DTT, 100 μM PMSF, pH 6.1 for six hours. This solution was then immediately chromatographed on a cation matrix, S-Sepharose (fast flow, obtainable from Pharmacia Corporation), pre-equilibrated in 20 mM potassium phosphate, 1 mM $MgCl_2$, 5 mM EGTA, 0.1 mM DTT, 100 μM PMSF, pH 6.1.

Proteins absorbed to the cation exchanger were eluted with a linear salt gradient containing 0-0.6 M sodium chloride. Using the GAP assay described below, most of the GAP activity was shown to be present in two peaks, a major peak eluting at a sodium chloride concentration of 100-150 mM, and a minor peak eluting at a sodium chloride concentration of 220-300 mM. The major peak was dialyzed against 30 mM Tris-HCl, 1 mM magnesium chloride, 1 mM EGTA, 0.1 mM DTT, 100 μM PMSF, pH 8.5. The dialyzate was applied to an anion exchange column, TSK-DEAE-5-PW (150 x 21.5 mm). The anion exchange matrix was treated with a linear salt gradient ranging from 0-0.6 M sodium chloride to elute the adherent proteins. Most of the GAP activity eluted at a sodium chloride concentration of about 130 mM

NaCl. Those fractions containing GAP activity were pooled, brought to 0.5 M ammonium sulfate, and passed through a hydrophobic column, phenyl-TSK HPLC. Proteins were eluted from the hydrophobic column using a crisscross gradient consisting of increasing ethylene glycol 0-30%, and decreasing ammonium sulfate, 0.5 M-0. The majority of GAP activity eluted at a concentration of 24% ethylene glycol and 0.1 molar ammonium sulfate. GAP activity assays, as performed below, correlated with a protein band of about 120,000 daltons, as revealed by sodium dodecyl sulfate polyacrylamide gel electrophoresis on 6% gels run under reducing conditions (Figure 1).

A second embodiment purification scheme was employed to purify GAP. Human placentas were again obtained shortly after delivery, and soaked in ice cold PBS, and homogenized and clarified as described in Example I. Ammonium sulfate was again added to the clarified homogenate to a final concentration of 40% to precipitate proteinaceous material. The ammonium sulfate solution was allowed to stand for one hour at 4°C prior to recovering the precipitated proteinaceous material by centrifugation for 15 minutes at 10,000 x g. The pellet was resuspended in PBS containing 0.1% NP40 and 100 μM PMSF. This solution was dialyzed for six hours at 4°C against 20 mM potassium phosphate, pH 6.1, containing 1 mM $MgCl_2$, 5 mM EGTA, 0.1 mM DTT, and 100 μM PMSF. Because GAP is susceptible to proteolysis, longer dialysis times are not desirable.

The GAP dialyzate was diluted three-fold with 4 mM potassium phosphate, pH 6.1, containing 0.02 M MgCl2, 1 mM EGTA, 0.1 mM DTT, and 100 μM PMSF to lower the conductivity of the solution to 1 millisiemens. This conductivity is compatible with application of the dialysate to a S-Sepharose cation exchange column. The dialysate was clarified by centrifugation at 10,000 x g for 10 minutes, followed by a further clarification step consisting of filtration through a 0.45 μM filter, prior to adding the dialysate to the S-Sepharose column (fast-flow, Pharmacia). Most of the contaminating proteins passed through the S-Sepharose column, and the adsorbed proteins eluted with a 1.5 liter salt gradient consisting of 0-0.6 M NaCl. Those fractions containing GAP activity were identified using the GAP assay described below.

As observed in the first example, GAP eluted from the cation exchange column in predominantly two major peaks. The first peak eluting over a sodium

chloride concentration of 100-150 mM was pooled and dialyzed against 30 mM Tris-HCl buffer, pH 8.5, containing 1 mM EGTA, 1 mM $MgCl_2$, 0.1 mM DTT and 100 μM PMSF. The solution was dialyzed at 4°C, and clarified by filtration with a 0.45 μM filter. The filtrate was divided into equal halves, and each half purified using two consecutive anion exchange columns.

The two filtrates were separately loaded onto a TSK-DEAE-5-EW column having the dimensions 150 x 21.5 mm. The column was pre-equilibrated in the Tris-hydrochloride, pH 8.5 dialysis buffer described above. GAP was eluted from the column with a 60-minute 0-0.6 M NaCl gradient with a flow rate of 3 ml/minute. The majority of the GAP activity from both filtrates eluted as a single peak at a sodium chloride concentration of about 130 mM. Sodium dodecyl sulfate, polyacrylamide gel electrophoretic analysis of the DEAE fractions showed that GAP was the major protein in the peak activity fractions. Fractions containing GAP from both purifications were pooled and diluted 5-fold into 2 mM potassium phosphate, pH 6.1, containing 0.1 mM EGTA, 10 μM DTT, 10 μM PMSF to lower the salt concentration to insure that the solution was chromatographically compatible with a second cation exchange chromatographic step, that is, chromatography with a SP-TFK column. The pH of the solution was checked and adjusted to pH 6.1 with sodium acetate (3 M, pH 4.8) if necessary. Both of the GAP fractions isolated from the DEAE columns were further purified separately over a cation column, TSK-SP-5-PW having dimensions of 75 x 7.5 mm. A solution containing 20 mM potassium phosphate, pH 6.1, containing 1 mM EGTA, 0.1 DTT, and 0.1 mM PMSF was passed through the column, followed by eluting GAP with a 45-minute, 0-0.6 M sodium chloride gradient at 1 ml per minute. Those fractions containing GAP were identified using the assay described below and sodium dodecyl sulfate polyacrylamide gel electrophoresis. GAP activity corresponded to a protein having a molecular weight of about 116,000 daltons. Amino acid analysis was performed on purified GAP to determine protein concentration. Starting with about 300 grams of human

placenta, approximately 430 micrograms of purified GAP was obtained. Figure 2 shows the SDS PAGE analysis of GAP at the various stages of purification described above.

GAP Assay

Several assays have recently been described to measure GAP activity. M. Trahey and F. McCormick, 1987 Science, 238:542; Adari et al., 1988 Science, 240:518. These references are herein incorporated in their entirety. GAP may be assayed in vitro, and several different types of in vitro assays can be performed. The preferred assay involves measuring the presence of GDP resulting from the hydrolysis of GTP. This assay involves combining in an appropriate physiologically buffered aqueous solution, empirically determined optimal amounts of normal cellular p21, and α-32P-GTP, plus GAP. The solution may also contain protease inhibitors and a reducing agent. Also, since cations greatly stimulate GAP activity they should be present in an effective amount. The preferred cation is magnesium chloride.

The reaction solution is incubated for various times and may be conducted at temperatures typically employed to perform enzymatic assays, preferably 10-40°C, and more preferably at 37°C. At the appropriate times aliquots are removed and assayed for α-32P-GDP. This is readily accomplished by first separating p21 containing bound α-32P-GDP from the other reactants in the solution, particularly free α-32P-GTP. This can be achieved by immunoprecipitating p21 with antibodies directed thereto. Immune precipitation techniques and anti-p21 antibodies are known, and routinely employed by those skilled in the art. α-32P-GDP, is released from the immune precipitate preferably by dissolving the sample in a denaturing detergent at an elevated temperature, more preferably in 1% sodium dodecyl sulfate at 65°C for five minutes, and chromatographing the mixture on a suitable thin layer chromatographic plate. The chromatography is preferably carried out on a PEI cellulose plate in 1 M LiCl. α-32P-GDP is identified by its mobility relative to a known standard using suitable radiodetection techniques, preferably autoradiography.

An alternative assay for GAP activity is to substitute gamma labeled 32P-GTP for a-labeled 32P-GTP in the above assay system, and assay for free 32P

labeled phosphate using activated charcoal. This assay can be carried out as described by Tjian et al., 1980 Cold Spring Harbor Symp. Quant. Biol., 44:103.

An additional assay does not involve immune precipitation. Rather, an aliquot from a GAP assay reaction mixture described above can be directly subjected to PEI cellulose chromatography in 1 M LiCl. This assay, however, is most useful for assaying solutions having substantially purified GAP.

A typical GAP assay can be carried out as follows. Approximately 0.8 micrograms of H-ras protein obtained as described by Trahey, et al., supra was bound to α-32P-GTP followed by precipitation of the complex with 13 micrograms of an anti-ras antibody, 157-181, that recognizes the carboxyl terminal end of the molecule. Specifically, 157-181 recognizes the carboxyl terminal residues at positions 157-181. Adari et al., 1988 Science, 280:518. Next, 10 micrograms of sheep-anti-mouse IgG, and 10 microliters of protein A-Sepharose beads were added. As a control, the same reactants were combined except that rat IgG replaced 157-181, and goat anti-rat IgG replaced sheep anti-mouse IgG. The pellets were washed with 20 mM tris hydrochloride, pH 7.4, containing 20 mM sodium chloride, 1 mM magnesium chloride and 1 mM DTT and resuspended in the same solution. Four microliter aliquots of the immune complex were then mixed with 10 microliters of GAP, or, as a control, buffer without GAP. After 60 minutes incubation at room temperature the Sepharose beads were washed again, and the bound nucleotides analyzed using thin layer chromatography with 1 M LiCl as the solvent. The thin layer plate was autoradiographed for one to two hours after which it was developed. The autoradiograph revealed that addition of sufficient GAP causes the near complete hydrolysis of GTP to GDP, whereas very little GTP hydrolysis occurs in the control lacking GAP. The assay detects GAP in a semi-quantitative, dose-dependent fashion. Quantitation can be improved by scraping the relevant regions of the plate and measuring cpm in GDP by use of a gamma counter. The immune precipitation controls having rat IgG substituted for the mouse antibodies revealed no GTP or GDP.

In addition to the above method, GAP can be preferably assayed as follows. Four μM normal cellular p21 was dissolved in a buffer containing 80 mM ß-glycerophosphate, 5 mM $MgCl_2$, 1 mM DTT, pH 7.5, plus 255 μM [α-32P] GTP (16 Ci/mmole), 4 mM ATP, and bovine serum albumin (2.5 mg/ml). The mixture

was preincubated for 30 minutes at 37°C, followed by the addition of either a sample suspected of containing GAP, or an equal volume of buffer. After one hour at room temperature the monoclonal antibody Y13-259 in the presence of 0.5% NP40 was added in an amount sufficient to bind all the p21 present in the solution. Next, goat anti-Rat Ig-Protein A Sepharose was added to collect p21 bound to Y13-259, and the immune complex isolated and washed ten times in 20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5 mM $MgCl_2$, and 0.5% NP40. To determine the extent of GTP binding and hydrolysis during these steps a control was run consisting of adding 5 μg of p21 immediately before adding Y13-259.

Nucleotides were eluted from p21 with 1% SDS, 20 mM EDTA at 65°C for five minutes and chromatographed on PEI Cellulose in 1 M LiCl. GTP and GDP were visualized using standard autoradiographic techniques. The results showed that normal cellular p21 affects a nearly complete conversion of GTP to GDP when compared to mutant ras oncogenic proteins Asp 12 and Val 12 assayed similarly. Moreover, little or no GTP or GDP was detected in the control sample.

The assays described above are presented in more detail by Trahey and McCormick, 1987 in Science, 238:542, and by Adari et al., 1988 in Science, 240:518. Both of these references are hereby incorporated by reference.

GAP Amino Acid Sequence

The GAP protein, or fragments derived therefrom can be sequenced using standard techniques known to those skilled in the art. In the event that GAP is isolated having a blocked amino terminal end, internal sequencing can be achieved by fragmenting the molecule such as, for example, with lysyl endopeptidase, and sequencing one or more of the resulting fragments. Although this may not necessarily be the case for GAP isolated from sources other than placenta, in the instant invention it was determined that GAP exhibited a blocked amino terminal end.

The protein having a molecular weight of about 120,000 obtained by the purification method described above was electro-eluted from a 6% sodium dodecyl sulfate, polyacrylamide gel in 0.05 molar ammonia bicarbonate containing 0.1% sodium dodecyl sulfate. The procedure followed is described by Hunkapillar et al., 1983 Methods in Enzymology, 91:227. The electro-eluted protein was fragmented for internal sequencing using lysyl endopeptidase (5% w/w, 18 hours at 40°C, WAKO).

Peptides were fractionated by reverse-phase high performance liquid chromatography using a Brownlee Aquapore RP-300 cartridge (100 x 2.1 mm, Applied Biosystems). Peptides were eluted with an Acetonitrile gradient from 0-70% in 120 minutes (Buffer A, 0.1% trifluoroacetic acid (TFA) in H20; Buffer B, 0.085% TFA in 85% acetonitrile). Automated sequence analysis of the peptides was conducted on an Applied Biosystems 470A gas-phase sequencer as reported. A peptide characteristic of GAP has the following amino acid sequence:

I M P E E E Y S E F K.

Cloning of GAP

A full length cDNA sequence that encodes GAP was obtained as follows: first, partial cDNA sequences were identified in a cDNA library using as oligonucleotide probes, DNA sequences derived from the partial amino acid composition of GAP. One such partial cDNA sequence, referred to as GAP 6, was subcloned and sequenced. Knowledge of its DNA sequence led, in turn, to additional probes that were used to screen cDNA libraries for longer cDNA inserts, eventually yielding the full length clone, clone 101. Each of the various procedures will be discussed below.

1. General Cloning Techniques

Construction of suitable vectors containing the desired GAP coding sequence employs standard ligation and restriction techniques which are well understood in the art. Isolated vectors, DNA sequences, or synthesized oligonucleotides are cleaved, tailored, and religated in the form desired.

Site specific DNA cleavage is performed by treating with suitable restriction enzyme(s) under conditions which are generally understood in the art, and the particulars of which are specified by the manufacturer of these commercially available restriction enzymes. See, e.g., New England Biolabs, Product Catalog. In general, about 1 µg of plasmid or DNA sequence is cleaved by one unit of enzyme in about 20 µl of buffer solution. In the examples herein, typically, an excess of restriction enzyme is used to insure complete digestion of the DNA substrate. Incubation times of about one hour to two hours at about 37°C are workable, although variations can be tolerated. After each incubation, protein is removed by extraction with phenol/chloroform, and may be followed by ether extraction, and the

nucleic acid recovered form aqueous fractions by precipitation with ethanol followed by chromatography using a Sephadex G-50 spin column. If desired, size separation of the cleaved fragments may be performed by polyacrylamide gel or agarose gel electrophoresis using standard techniques. A general description of size separations is found in Methods in Enzymology (1980) 65:499-560.

Restriction cleaved fragments may be blunt ended by treating with the large fragment of E. coli DNA polymerase I, that is, the Klenow fragment, in the presence of the four deoxynucleotide triphosphates (dNTPs) using incubation times of about 15 to 25 minutes at 20 to 25°C in 50 mM Tris pH 7.6, 50 mM NaCl, 6 mM $MgCl_2$, 6 mM DTT and 10 mM dNTPs. After treatment with Klenow, the mixture is extracted with phenol/chloroform and ethanol precipitated. Treatment under appropriate conditions with S1 nuclease results in hydrolysis of single-stranded portions.

Ligations are performed in 15-30 µl volumes under the following standard conditions and temperatures: 20mM Tris-Cl pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 33 µg/ml BSA, 10 mM-50 mM NaCl, and 1 mM ATP, 0.3-0.6 (Weiss) units T4 DNA ligase at 14°C for "sticky end" ligation, or for "blunt end" ligations 1 mM ATP was used, and 0.3-0.6 (Weiss) units T4 ligase. Intermolecular "sticky end" ligations are usually performed at 33-100 µg/ml total DNA concentration. In blunt end ligations, the total DNA concentration of the ends is about 1 µM.

In vector construction employing "vector fragments," the vector fragment is commonly treated with bacterial alkaline phosphatase (BAP) in order to remove the 5' phosphate and prevent religation of the vector. BAP digestions are conducted at pH 8 in approximately 150 mM Tris, in the presence of Na^+ and Mg^{+2} using about 1 unit of BAP per µg of vector at 60°C for about 1 hour. Nucleic acid fragments are recovered by extracting the preparation with phenol/chloroform, followed by ethanol precipitation. Alternatively, religation can be prevented in vectors which have been double digested by additional restriction enzyme digestion of the unwanted fragments.

In the constructions set forth below, correct ligations are confirmed by first transforming the appropriate E. coli strain with the ligation mixture. Successful transformants are selected by resistance to ampicillin, tetracycline or other antibiotics, or using other markers depending on the mode of plasmid construction, as is understood in the art. Miniprep DNA can be prepared from the transformants by the method of D. Ish-Howowicz et al., (1981 Nucleic Acids Res. 9:2989) and analyzed by restriction and/or sequenced by the dideoxy method of F. Sanger et al., 1977 Proc. Natl. Acad. Sci. (USA), 74:5463 as further described by Messing et al., 1981 Nucleic Acids Res., 9:309, or by the method of Maxam et al., 1980 Methods in Enzymology, 65:499.

Host strains used in cloning in M13 consists of E. coli strains susceptible to phage infection, such as E. coli K12 strain DG98 are employed. The DG98 strain has been deposited with ATCC July 13, 1984 and has accession number 1965.

Depending on the host cell used, transformation is done using standard techniques appropriate to such cells. The calcium treatment employing chloride, as described by S. N. Cohen, 1972 Proc. Natl. Acad. Sci. (USA) 69:2110, or the $RbCl_2$ method described in Maniatis et al., 1982 Molecular Cloning: A Laboratory Manual Cold Spring Harbor Press, p. 254 was used for procaryotes. Transfection of Sf9 cells was achieved using a modification of the calcium phosphate precipitation technique (Graham, F.L. et al., 1973 Virology 52:456) as adapted for insect cells (J. P. Burand et al., 1980 Virology 101; E. B. Casstens et al., 1980 Virology 101:311). Additional details regarding transfection of Sf9 cells are described by Summers and Smith in "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures", Texas A & M Press: 1986. The baculovirus transfer vectors employed herein are derived from transfer vectors which have been described by G. E. Smith et al., 1983, above. These vectors were originally constructed by cloning the AcNPV EcoRI-1 fragment containing the polyhedrin gene into the Eco RI site of E. coli plasmid pUC8 as described by Vieira et al., 1982 Gene 19:259-268. A family of plasmids having single Bam HI cloning sites at various locations in the polyhedrin gene were created as described by Smith et al., 1983, above. The most used of these, pAc373, has a

unique Bam HI site 50 base pairs downstream from the polyhedrin cap site, that is to say, 8 base pairs before the polyhedrin ATG translation initiation codon (Luckow and Summers in Biotechnology, Vol. 6, p. 47 (1988).

2. Oligonucleotide Probes

Synthetic oligonucleotides were prepared by the triester method of Matteucci et al., 1981 J. Am Chem. Soc. 103:3185 or using commercially available automated oligonucleotide synthesizers. Kinasing of single strands prior to annealing or for labeling is achieved using an excess, e.g., approximately 10 units of polynucleotide kinase to 0.1 nmole substrate in the presence of 50 mM Tris, pH 7.6, 10 mM $MgCl_2$, 5mM dithiothreitol, 1-2 mM ATP, 1.7 pmoles gamma ^{32}P-ATP (2.9 mCi/mmole), 0.1 mM spermidine, 0.1 mM EDTA.

Using the partial GAP amino acid sequence described above, and known codon redundancies thereto, several DNA oligonucleotide probes were synthesized and these are shown in Figures 3 and 4.

3. Identification and Isolation of GAP Sequences

Several procedures are available for identifying GAP DNA sequences. The preferred procedure is to use the oligonucleotide probes described above to screen cDNA libraries. cDNA libraries can be constructed using techniques known in the art, or can be purchased commercially.

An illustrative procedure for making a cDNA library containing GAP sequences may consist of isolating total cytoplasmic RNA from suitable starting material, and further isolating messenger RNA therefrom. The latter can be further fractionated into Poly (A+) messenger RNA, which in turn is fractionated further still into Poly (A+) messenger RNA fractions containing GAP messenger RNA. The appropriate GAP messenger RNA can then be reverse transcribed and cloned into a suitable vector to form the cDNA library.

More specifically, the starting material (i.e., tissue, cells) is washed with phosphate buffered saline, and a non-ionic detergent, such as ethylene oxide, polymer type (NP-40) is added in an amount to lyse the cellular, but not nuclear membranes, generally about 0.3%. Nuclei can then be removed by centrifugation at 1,000 x g for 10 minutes. The post-nuclear supernatant is added to an equal volume of TE (10

mM Tris, 1 mM ethylenediaminetetraacetic acid (EDTA), pH 7.5) saturated phenol/chloroform (1:1) containing 0.5% sodium dodecyl sulfate (SDS) and 10 mM EDTA. The supernatant is re-extracted 4 times and phase separated by centrifugation at 2,000 x g for 120 minutes. The RNA is precipitated by adjusting the samples to 0.25 M NaCl, adding 2 volumes of 100% ethanol and storing at -20°C. The RNA is then pelleted at 5,000 x g for 30 minutes, washed with 70% and 100% ethanol, and dried. This represents the total cytoplasmic RNA. Polyadenylated (Poly A+) messenger RNA (mRNA) can be obtained from the total cytoplasmic RNA by chromatography on oligo (dT) cellulose (J. Aviv et al., 1972 Proc. Natl. Acad. Sci. 69:1408-1412). The RNA is dissolved in ETS (10 mM Tris, 1 mM EDTA, 0.5% SDS, pH 7.5) at a concentration of 2 mg/ml. This solution is heated to 65°C for 5 minutes, then quickly chilled to 4°C. After bringing the RNA solution to room temperature, it is adjusted to 0.4 M NaCl and slowly passed through an oligo (dT) cellulose column previously equilibrated with binding buffer (500 mM NaCl, 10 mM Tris, 1 mM EDTA, pH 7.5) The flow-through is passed over the column twice more, and the column washed with 10 volumes of binding buffer. Poly (A+) mRNA is eluted with aliquots of ETS, extracted once with TE-saturated phenol chloroform and precipitated by the addition of NaCl to 0.2 M and 2 volumes of 100% ethanol. The RNA is reprecipitated twice, washed once in 70% and then 100% ethanol prior to drying. The poly (A+) mRNA can then be used to construct a cDNA library.

cDNA can be made from the enriched mRNA fraction using oligo (dT) priming of the poly A tails and AMV reverse transcriptase employing the method of H. Okayama et al., 1983 Mol. Cell Biol. 3:280, incorporated herein by reference.

Other methods of preparing cDNA libraries are, of course, well known in the art. One, now classical, method uses oligo (dT) primer, reverse transcriptase, tailing of the double stranded cDNA with poly (dG) and annealing into a suitable vector, such as pBR322 or a derivative thereof, which has been cleaved at the desired restriction site and tailed with poly (dC). A detailed description of this alternate method is found, for example, in U.S. Serial No. 564,224, filed December 20, 1983, and assigned to the same assignee, incorporated herein by reference.

As mentioned above, cDNA libraries are commercially available. A particularly useful library is sold by Clontech (Catalog number #L H1008). It is a lambda gt11 human placenta cDNA library made from total poly (A+) messenger RNA.

4. Identification of GAP DNA Sequences

The oligonucleotide probes described above, GW13, GW15, GW17 and GW19 were used to screen the commercially available Clontech library. The library was plated at about 50,000 plaques per plate using 17 plates. Thus, about 850,000 plaques were screened using the plaque hybridization procedure. While a variety of such procedures are known, a description of the preferred procedure follows. Each 150 mM plate was replicated onto duplicate nitrocellulose filter papers (S & S type BA-85). DNA was fixed to the filter by sequential treatment for 5 minutes with 0.5 N NaOH plus 1.0 M NaCl; 1.5 M NaCl plus 0.5 M Tris-HCl pH 8; and 20mM Tris plus 2mM EDTA pH 8. Filters were air dried and baked at 80°C for 2 hours.

The duplicate filters were prehybridized at 55°C for 2 hours with 10 ml per filter of DNA hybridization buffer, 5x SSC, pH 7.0, 5x Denhardt's solution (polyvinylpyrrolidone, plus Ficoll and bovine serum albumin; 1 x 0.02% of each), 50 mM sodium phosphate buffer at pH 7.0, 5 mM EDTA, 0.1% SDS, and 100 µg/ml yeast RNA. The prehybridization buffer was removed and the samples were hybridized with a mixture of kinased probes under conditions which depend on the stringency desired. About 2 x 10^6 cpm/ml total was used. Typical moderately stringent conditions employ a temperature of 42°C plus 50% formamide for 24-36 hours with 1-5 ml/filter of DNA hybridization buffer containing probe. For higher stringencies high temperatures and shorter times were employed. The preferred hybridization conditions consisted of hybridizing the probes to the filters in 5 x SSC (standard saline citrate), Denhardt's solution, 50 mM $NaPO_4$ pH 7.0, 5 mM EDTA, 0.1% SDS, and 100 mg/ml yeast RNA at 55°C overnight. Next, the filters were washed twice, 30 minutes each wash, at room temperature with 2 x SSC, 0.1% SDS and 50 mM sodium phosphate buffer pH 7, then washed once with 2 x SSC and 0.1% SDS at 50°C, and air dried. Finally, the filters were autoradiographed at -70°C for 36 hours.

The autoradiographic results revealed a single positive plaque. Using the

washing and hybridization conditions described above, several lambda gt11 plaque purified isolates were identified and picked. Viral DNA was obtained from one of these, termed GAP 6, as follows. GAP 6 was plated at high density on a lawn of E. coli strain Y 1090 (r^-). Following lysis of the E. coli, phage particles were eluted into S M buffer (0.1 M NaCl 8.1 mM $MgSO_4$ 50 mM Tris - HCl pH 7.5 0.01% Gelatin) by covering the E. coli with buffer and incubating the plate in the cold for several hours. The lysate containing phage particles was centrifuged at 11,500 xg for 20 minutes to remove cellular debris, and the resulting supernatant titered using standard techniques. A titer of 2 x 10^10 PFU/ml was determined. Finally, phage DNA was isolated by the procedure of Maniatis et al., above.

5. Characterization of GAP 6

GAP 6 was subcloned into a suitable vector in order to characterize the DNA both as to Eco RI restriction sites, and partial DNA sequence. Although GAP 6 DNA can be cloned into a variety of vectors, in the instant invention it was cloned into M13. More specifically GAP DNA was cloned into a M13 vector as follows. GAP 6 DNA was treated with Eco RI enzyme which produced two fragments, about 2.0 kb and 0.24 kb. These fragments were isolated using standard agarose gel techniques, and ligated into M13mp18. The M13 vector was designed so that vectors without DNA inserts show up blue under the proper culture conditions, whereas vectors with a DNA insert are clear.

The ligated M13mp18 phage were transduced into frozen competent E. coli K12 strain DG98 and cultured by plating on media containing 5 x 10^{-4} M isopropyl thiogalactoside (IPTG) obtained from Sigma Chem. (St. Louis, MO) and 40 µg/ml X-gal. Non alpha-complementing white plaques were picked onto fresh media. Mini-cultures were screened for recombinant single strand phage DNA containing inserts.

The white M13 plaques were screened for inserts by direct gel electrophoresis. The latter procedure was conducted essentially as described by J. Messing, 1983 Methods of Enzymology 101:20, which is hereby incorporated by reference. Four M13mp18 subclones were identified by this method. Two subclones, GAP 2 and GAP 8, contained the 2kb fragment in both orientations. The remaining two subclones, GAP 12 and GAP 18, contained the 0.24 kb fragment in both orientations.

The partial DNA sequence of GAP 2 and GAP 8 was determined by the T. Sanger, S. Nicklen, and H.R. Coulson, 1977 Proc. Natl. Acad. Sci. USA 74:5463-5467 techniques described above:

5' AAAACTCATGC AAGGGAAGGG CAAAACCCAG TATGGTCAGA
AGAGTTTGTC TTTGATGATC TTCCTCCTGA CATCAATAGA
TTTGAAATAA CTCTTAGTAA TAAAACAAAG AAAAGCAAAG
ATCCTGATAT CTTATTTATG CGCTGCCAGT TGAGCCGATT
ACAGAAAGGG CATGCCACAG ATGAATGGTT TCTGCTCAGC
TCCCATATAC CATTAAAAGG TATTGAACCA GGGTCCCTGC
GTGTTCGAGC ACGATACTCT ATGGAAAAAA TCATGCCAGA
AGAAGAGTAC AGTGAATTTA AAGAGCTTAT ACTGCAAAAG
GAACTTCATG TAGTCTATGC TTTATCACAT 3'

6. Identification of GAP DNA Sequences Longer Than GAP 6

General Technique: A novel procedure was used to identify plaques that contain GAP cDNA inserts larger than those present in GAP 6 which consisted of elucidating inserts present in either the lambda gt11 library described above, or a lambda gt10 library described below. The procedure consisted of synthesizing cDNA inserts using DNA oligonucleotides having sequences complementary to the 5' region of GAP 6, and oligonucleotide primers that flank the EcoRI insertion sites of lambda gt11, or lambda gt10, using the polymerase chain reaction, or PCR. The newly identified PCR products were sequenced, and accordingly DNA probes were synthesized having sequences 5' of GAP 6. These probes were, in turn, used to identify plaques containing larger GAP cDNA inserts. The procedure was repeated several times using as probes, DNA sequences progressively further 5' of GAP 6 identified from each round of newly synthesized cDNA inserts.

PCR is described in U.S. Patents 4,683,202 and 4,683,195, both of which are hereby incorporated in their entirety. In general, the synthesis/amplification of DNA sequences by PCR involves an enzymatic chain reaction that produces, in exponential quantities, a specific DNA sequence, provided that the termini of the sequence are known in sufficient detail so that oligonucleotide primers can be synthesized which will hybridize to them, and that a portion of the sequence is available to initiate the chain reaction. One primer is complementary to the negative strand, and the other is

complementary to the positive strand. As applied to the instant invention, the primers employed are complementary to the 5'end of GAP 6, and are complementary to and flank the Eco RI sites of lambda gt11, or lambda gt10. Because the orientation of a particular cDNA insert in either vector is not known, it was necessary to run separate reactions with oligonucleotides that flank both sides of the Eco RI site. Exemplary of primers useable with lambda gt11 are two 24-base sequencing primers, 1218 and 1222, produced by New England Biolabs. Similarly, primers compatible with lambda gt10 are also available from New England Biolabs, and these are 1231 and 1232. Thus, separate reactions were run with either 1218, 1219, or 1231 and 1232, and the appropriate GAP 6 primer.

The primers are annealed to denatured DNA acid, followed by extension with a suitable DNA polymerase enzyme, such as the large fragment of DNA polymerase I (Klenow), or preferably a DNA polymerase that is stable in the presence of detergents and nucleotides, which results in newly synthesized plus and minus strands containing the target sequence. Alternatively, a thermostable enzyme may be used which is present in thermostable bacteria. The enzyme may be produced using DNA recombinant techniques as described in U.S. Patent Application Serial No. 063,509, Filed July 17, 1987. Because the newly synthesized sequences are also templates for the primers, repeated cycles of denaturing, primer annealing and extension results in exponential accumulation of the region defined by the primer. PCR thus produces discrete nucleic acid duplexes of cDNA inserts having termini corresponding to the ends of the specific primers employed.

Although PCR can be performed using a variety of reaction conditions, as described in the references presented above, the preferred reaction conditions are as follows. Plaques that hybridize to a particular probe are eluted into either 0.5ml of water, or SM buffer, and 50 µl of the eluate combined with 10 µl of 10 x PCR buffer, 1.5µl of 10 mM dNTP's, 1µl of a first and second primer, each at a

concentration of about 20 pmoles, 0.2 µl of Taq polymerase equivalent to 1 unit of activity. The final volume is 100 µl. PCR 10 x buffer consists of 500 mM KCl, 200 mM Tris-HCl, pH 8.4, 25 mM $MgCl_2$ and 1 mg/ml.

GAP encoding sequences: Gap 6 DNA was sequenced, and an oligonucleotide probe based on the sequence, GW50, synthesized, radiolabelled, and used to rescreen the Clontech lambda gt11 library, and to screen a second cDNA library made from K562 cells. K562 cDNA was cloned in lambda gt10, and a description of this library is presented by Mes-Masson et al., 1986, in the Proceedings of the National Academy of Sciences, 83, 9768. This publication is hereby incorporated by reference in its entirety. The oligonucleotide, GW50, has the following sequence:

5' TTTAAATTCACTGTACTCTTCTTCTGGCATGAT 3'

Hybridization of GW50 to either library was conducted as described above with the exception that the washing steps after the hybridization were more stringent. Specifically, the filters containing plaques were washed twice, for 15 minutes each wash, with 2 x SSC containing 0.1% SDS at room temperature and then two additional washes, for 15 minutes each, with 0.2 x SSC containing 0.1% sodium dodecyl sulfate at 55°C. Autoradiography of the filters prepared from the Clontech library revealed 160 positive plaques, while only one plaque was detected from the K562 library.

Using the sequence of GAP 6, DNA primers, LC121 and LC122, with sequences complementary to the 5' region of GAP 6, were synthesized.

LC121 5' GAGGAAGATCATCAAAGACAAACTCT 3'

LC122 5' TCTGTAATCGGCTCAACTGGCAGCG 3'

LC121 corresponds to the 5' end of GAP 6 in the anti-sense direction.

The 163 positive plaques from the Clontech library, and the one positive plaque from the K562 library, were removed from agarose plates using a Pasteur pipette, and eluted into 0.5ml of SM buffer for 30 minutes. Each isolate was then PCR processed as described above using LC121 in combination with the appropriate lambda primers. Typically, a denaturation step was run for 2 minutes at 94°C,

followed by an annealing step for 30 seconds at 55°C, and an extension step for 5 minutes at 72°C. The reaction was most often run for 30 cycles. The resulting amplified cDNA inserts were sequenced.

Sequencing can be performed using the techniques referred to above, or by direct sequencing of single stranded DNA produced by PCR. The use of PCR to generate single stranded DNA is described in a co-pending U.S. Patent Application, Serial No. 248,896, titled "Method for Generating Single Stranded DNA by the Polymerase Chain Reaction", Filed on September 23, 1988. This patent application is hereby incorporated by reference in its entirety.

Typically about 50 µl of the PCR reaction was separated on a 1% agarose TAE gel, the region of the gel containing the amplified products excised, and the PCR products extracted from the agarose and suspended in about 10 µl-20 µl of TE buffer. Generally about one tenth of this volume was subjected to asymmetric PCR amplification. The reaction conditions employed are described in the above cited patent application. The primers were typically used in a ratio of about 100:1, or about 50:0.5 pmoles.

Using LC121, 14 of the 163 lambda gt11 plaques were found to have an additional 320, or greater number of base pairs 5' of GAP 6, while the single plaque isolated from the K562 lambda gt10 library, referred to as K16, was determined to have a cDNA insert consisting of GAP 6 plus an additional 700 base pairs 5' thereto. Based on the latter sequence, several additional oligonucleotides, LC136, LC138, and LC140 were synthesized and used in conjunction with LC121 to again screen the 163 plaques from the Clontech library. The primers have the following sequences:

LC136 5' CGTAAATTGCAAAATGCCTGCAGACCTTG 3'

LC138 5' GTTTTCCTTTGCCCTTTTTCAGAAGATAAC 3'

LC140 5' TGTCATTGAGTACTTGTTCTTGATCCTGC 3'

Rescreening the 163 plaques with LC136 revealed that 82 plaques were positive, while rescreening with LC138 plus LC140, revealed that 63 of the plaques were positive. Of the 63 positive plaques, 38 were subjected to PCR using the primers 1218 and LC138; and 1222 and LC138. Of these, six were found to have long stretches of DNA 5' to GAP 6. Sequencing in M13m18 revealed that they represent different length fragments of the same type of transcript. Two of the clones were studied in detail, clone 7 and clone 101. Clone 101 contained sufficient DNA

to encode a protein of 1047 amino acids, which would have a molecular weight of 116,000 daltons. This is similar to the molecular weight of the GAP protein purified from human placenta as described above. Thus, clone 101 contains the full length GAP cDNA. Clone 101 was sequenced, and the sequence is shown in Figure 5. Clone 7 was also sequenced, and shown to have the identical sequence as clone 101 but lacking 33 base pairs from the 5' end.

In addition to the above, two plaques were identified from the 163 plaques initially determined to be positive with GW50 that contained cDNA inserts consisting of an additional 65 base pairs inserted between nucleotides 537 and 538 of clone 101. One of the two clones, clone 16, lacks the first 180 amino acids of clone 101, while the other clone, clone "Sleepy", lacks at least the first 180 amino acids, and additionally has a truncated 3' end at about base 2448 of clone 101. The DNA sequence of the 65 base pair insert is shown in Figure 6 for clone 16.

7. Expression of GAP

Lambda lysogens: GAP activity was detected from lysates of lambda lysogens of clones 7, 16, and 101. Lysogens were generated in E. coli strain Y1089. The procedures for growing, inducing, harvesting, and lysing the cells is described in T. Huynh et al., in "DNA Cloning Techniques: A Practical Approach", D. Glover, Ed. (IRL press, Oxford, 1985) pp 49-78. This publication is hereby incorporated by reference in its entirety. Briefly, supernatants obtained from lysates were dialyzed into GAP assay buffer consisting of 20 mM Tris-HCl, pH 7.0, 1 mM $MgCl_2$, 0.1 mM DTT, 0.1% NP40, 100 μM PMSF, and GAP activity measured using the TLC-based GTPase assay described above. 2.2μM of either normal N-ras p21 protein having glycine at position 12, or mutant p21 proteins wherein glycine is substituted with aspartic acid or valine, were incubated with 0.25μM [α-^{32}P] GTP (800Ci/mmole) for 15 minutes at 37°C in the presence or absence of lambda lysate. As discussed earlier, the mutant p21 proteins have transforming activity and do not exhibit significant GAP stimulatable GTPase activity. About 10 μl of lysate or GAP assay buffer was added, and after 1 hour at room temperature, p21 was immunoprecipitated and associated nucleotides analyzed by chromatography on PEI cellulose in 1 M LiCl. An additional control was run for GAP activity; it consisted of testing an irrelevant lysogen lysate, specifically lambda gt11 lacking a cDNA insert. The results

are shown in Figure 7 for clones 7 and 101. The upper part of panel A shows the results for clone 7, while the lower region of the panel shows the results for clone 101. It is apparent that lysates from both clones stimulate the hydrolysis of GTP to GDP in the presence of normal p21, but not in the presence of mutant p21 proteins. Moreover, when GAP buffer is substituted for normal p21, or the mutants, there was no effect on GTP hydrolysis. The irrelevant lysogen lysate also did not support GTP hydrolysis.

Transfection of Spodoptera frugiperda: The full length cDNA insert in clone 101 was expressed in insect cells, Spodoptera frugiperda. The insect cell line, Sf9, was transfected with a baculovirus expression vector, pAcC12, containing the GAP encoding Eco R1 fragment of clone 101, and GAP activity measured in cell extracts.

The baculovirus vector, pAcC12, was constructed from preexisting vectors, particularly pAc311 and pAc373, as described by Luckow and Summers in Biotechnology, Vol. 6, p. 47 (1988); U.S. Patent No. 4,745,051; and EPA 127,839. Additional details are presented by Summers and Smith in "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures", Texas Agricultural Experiment Station Bulletin No. 1555, May, 1987. All of these references are hereby incorporated in their entirety.

pAcC12 was constructed as described below, and as shown in Figure 8. The transfer vector pAc311 was site-directed mutagenized using M13 mutagenesis techniques to convert the polyhedrin gene initiation codon, ATG, to ATT. The resulting vector was designated pVL941, and is described in detail by Luckow and Summers in Virology, titled "High Level of Expression of Non-Fused Foreign Genes with Autographa Californica Nuclear Polyhedrosis Virus Expression Vectors". A polylinker was inserted into pVL941 at a unique BamHI site 30 base pairs downstream of the ATT sequence. pVL941 was digested with Bam HI, and the polylinker, consisting of two complementary self-annealed oligomers, EK 129 and EK130, having the sequences shown below, ligated to produce the vectors pAcC8 and

pAcC9 that carry the polylinker in different orientations. The polylinker has a restriction site for Eco RI, as well as for other restriction enzymes.

EK 129:

5'GATCCACCATGGAGCTCGAGATCTAGAATTCTGCAGCCCGGGTACCGATC 3'

EK 130:

5'GATCGGTACCCGGGCTGCAGAATTCTAGATCTCGAGCTCCATGGTGGATC 3'

Because pAcC8 and pAcC9 have two Eco RI restriction sites, one in the polylinker and the other in the plasmid DNA as shown in Figure 8, it was desirable to remove the plasmid Eco RI site so that the Gap Eco RI encoding fragment of clone 101 could be inserted into the polylinker site. This was achieved using the transfer vector pAc373. pAc373 is similar to pAc311 except that the nucleotide sequences spanning the polyhedrin start codon differ. Thus, the Eco RI site was removed from pAc373 by digesting the vector to completion with Eco RI, and the ends made blunt using the Klenow fragment under the appropriate reaction conditions. Following ligation and transformation into E. coli DH 5, colonies were identified that lacked the Eco RI site by restriction analysis of miniprep DNA.

pAc373 lacking the Eco RI site was further modified by incorporating the polylinker consisting of the oligomers, EK129 and EK130, shown above, by digesting the vector with Bam HI, followed by ligating the oligomers. The resulting vectors, pAcC6 and pAcC7, contain the polylinker in different orientations.

The final construct, pAcC12, was generated from pAcC7 and pAcC9 as shown in Figure 8. These vectors contain the polylinker in the same orientation. Both vectors were digested with Bst EII and Eco RI and the resulting fragments electrophoretically purified. The Bst EII/Eco RI fragment of pAcC7 containing the pUC 8 sequences, and partial polylinker sequences was ligated to the large BstEII/Eco RI fragment of pAcC9. This latter fragment contains the ATT sequence and the remaining polylinker sequences.

The transfer vector, pAcC12, has the Eco RI GAP fragment of clone 101 inserted in both orientations. The correct orientation was designated pAcC12 GAP 5, while the incorrect orientation was designated pAcC12GAP 101-7. About 2μg of either plasmid was transfected into 2 x 10^5 Sf9 cells, the cells grown for 4 days, isolated by centrifugation, and cell extracts made by solubilizing the cell pellet. The preferred solubilization solution consists of 0.5% NP40, 10 mM Tris HCl, pH 8.0,

and 150 mM NaCl. The extract was centrifuged for 15 minutes at 15,000xg and aliquotes diluted into GAP assay buffer, and assayed for GAP activity as described above. Methods for growing Sf9 cells are well known in the art, and detailed procedures for their cultivation can be found in M. Summers and G. Smith in "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures", Texas Agricultural Experiment Station, Bulletin No. 1555 (May, 1987) or in EPO 127,839 to G.E. Smith and M. D.Summers. Preferred media and culturing conditions can be found in co-pending, commonly owned U.S. Patent Application, Serial Nos. 77,181, titled "Airlift Insect Cell Culture, Filed July 24, 1987; 77,303, titled "Serum Free Media for the Growth of Insect Cells and Expression of Products Thereby", and 77,189, titled "Lipid Microemulsions for Culture Media". These publications and patent applications are hereby incorporated by reference. Figure 7, panel B shows the results. The effect of pAcGAP 5 and pAcGAP 101-7 are shown in lanes 1 and 2, respectively; lane 3 presents a buffer control. Note that pAcGAP 5- stimulates normal ras p21 GTPase activity, whereas it is without effect on the p21 mutants. In contrast, there is no stimulation of GTPase activity by pAcGAP 101-7 of either normal ras p21 or the mutants.

It is important to note that baculovirus can be recovered from Sf9 cells transfected with the above described transfer vectors using the techniques described by Summers and Smith, above. Such virus can be employed to transform cells directly with the appropriate GAP clone.

8. Diagnostic Uses of GAP Sequences

The GAP DNA sequences described herein can be used to produce GAP, which, in turn, can be used to produce antibodies to GAP. These antibodies may be used to isolate GAP using antibody purification techniques generally known in the art. Since GAP is one reagent employed in assaying for the presence of normal ras p21, as described above, especially in tumors thought to result from the overexpression of ras p21, and is now available only in limited amounts because of the burdensome purification methods used to obtain it, the availability of large amounts of GAP will be a valuable addition to present cancer diagnostic methods.

The GAP DNA sequences disclosed herein may also be used to determine the number of copies of the GAP gene present per cell in various types of cancers,

that is to say, whether the gene is amplified. It is applicant's belief that tumors thought to be causally related to ras expression, overexpress GAP via gene amplification. Thus, the GAP DNA sequences disclosed herein can be used to measure the degree of overamplification, and diagnostic and prognostic correlations established.

The GAP DNA sequences can be used to measure the level of amplification following techniques generally known in the art. D. Slamon et al., 1987 Science 235:177; U.S. Patent 4,542,092 and U.S. Patent 4,699,877; R. Schimke, 1982 Gene Amplification, Cold Spring Harbor Laboratory. These publications are hereby incorporated by reference in their entirety. GAP gene amplification can be measured directly using established DNA hybridization techniques. DNA is prepared from human tumor tissue as described by Maniatis et al., and Slamon et al., above, or J. Southern, 1975 Mol. Biol. 98:503, and reacted with labeled GAP DNA. GAP 6, GAP 2 or GAP 8 sequences may be used. The entire sequence may be used, or short nucleotide sequences derived therefrom. Normally, a sequence should have at least about 14 nucleotides, and preferably at least about 18 nucleotides. Various labels may be employed to label the GAP sequences, most commonly radionuclides, particularly 32 P are used. However, other techniques may also be employed, such as using biotin modified nucleotides for introduction into a polynucleotide. The biotin then serves as the site for binding to avidin or antibodies, which may be labeled with a wide variety of labels, such as radionuclides, fluorescent molecules, enzymes, or the like. Alternatively, antibodies may be employed which can recognize specific duplexes, including DNA duplexes, RNA duplexes and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected. The DNA probe labelling procedures are known in the art. See, Maniatis et al., above.

A suitable DNA preparation and hybridization procedure to determine the number of GAP genes per cell using dilutional analysis as described by Slamon et al., above, consists of extracting and digesting tumor DNA using the procedure of Maniatis et al., above, followed by subjecting about 10-15 μg of EcoRI digested DNA to electrophoresis in 0.8% agarose gel, or dilutions of the digested DNA and transferring the DNA onto nylon filter papers. The filters are baked in a vacuum oven for 2 hours at 80°C, prehybridized in 5 x SSC solution containing 50% formamide, 10% dextran sulfate, 0.1% SDS, denatured salmon sperm DNA (1mg/ml), and 4 x Denhardt's solution for 12 hours. The DNA can then be hybridized in the same solution containing 32 P-labeled nick-translated GAP 8 probe with a specific activity of about 1 x 10^8 cpm/μg DNA, or about 2 x 10^6 cpm/ml. Optimal hybridization and washing conditions can be especially determined, however results may be apparent if hybridization occurs for 48 hours at 42°C, and the filters are washed in succession as follows: 2 x SSC for 20 minutes at room temperature; two washes of 30 minutes each in 2 x SSC, 0.1% SDS at 65°C; and one wash of 30 minutes in 0.5 x SSC, 0.1% SDS at 65°C. Filters can then be exposed to x-ray film for autoradiography, and the degree of amplification ascertained using established methods, including soft laser densitometry scanning.

Using the above techniques, a correlation may be observed wherein individuals with tumors that have 2-4 copies of GAP enjoy a favorable diagnosis and are unlikely to develop an aggressive malignancy, whereas tumors with 4 or more copies are likely to have aggressive malignancies, and require extensive medical treatment.

In addition to directly detecting GAP gene amplification by the foregoing procedures, amplification may also be detected indirectly by measuring GAP gene messenger RNA levels with labelled GAP DNA sequences. The procedures generally applicable to this method are described in Maniatis et al., above, and in U.S. Patent 4,699,877.

Deposit of Biological Materials: The following plasmids have been deposited with the American Type Culture Collection on October 11, 1988.

Designation	ATCC No.	CMCC No.
pAcC GAP 5 (pAcC12 GAP 5)	67821	3437
pGAP 16-4 (Clone 16)	40503	3479
pGAP-SLE1 (Clone Sleepy)	40504	3480

Having generally described the invention, it will be appreciated that the scope of the invention is limited only by the appended claims, and not by the particular materials and methods described above.

WE CLAIM:

1. A recombinant DNA molecule encoding a polypeptide possessing GAP activity comprising the sequence shown in FIG. 5.

2. A recombinant DNA molecule described in Claim 1 wherein said polypeptide encoded by said molecule has a blocked amino terminal end.

3. A recombinant DNA molecule as described in Claim 1, encoding a polypeptide possessing GAP activity further comprising a 65 base pair insert as shown in FIG. 6.

4. A recombinant DNA molecule encoding a polypeptide possessing GAP activity, and having the sequence shown in FIG. 5, operably linked to appropriate control sequences which regulate the synthesis of said recombinant DNA molecule and the expression of said polypeptide in transformed host cells.

5. The recombinant DNA of Claim 4 wherein said host cells are procaryotic or eucaryotic cells.

6. The recombinant DNA of Claim 5 wherein said procaryotic host cells are Escherichia coli.

7. The recombinant DNA of Claim 5 wherein said eucaryotic cells are insect cells.

8. The recombinant DNA of Claim 7 wherein said eucaryotic cells are insect cells.

9. The recombinant DNA of Claim 8 wherein said insect cells are Spodoptera frugiperda cells.

10. A recombinant DNA molecule encoding a polypeptide possessing GAP activity, and having the sequence shown in FIG. 6, operably linked to appropriate control sequences which regulate the synthesis of said recombinant DNA molecule and the expression of said polypeptide in transformed host cells.

11. The recombinant DNA of Claim 10 wherein said host cells are procaryotic or eucaryotic cells.

12. The recombinant DNA of Claim 11 wherein said procaryotic host cells are Escherichia coli.

13. The recombinant DNA of Claim 11 wherein said eucaryotic cells are insect cells.

14. The recombinant DNA of Claim 13 wherein said eucaryotic cells are insect cells.

15. The recombinant DNA of Claim 14 wherein said insect cells are Spodoptera frugiperda cells.

16. Baculovirus transfer vectors comprising the DNA sequence of Claim 1.

17. The baculovirus transfer vectors of Claim 16 wherein said vectors are selected from the group consisting of pAcC12-GAP 5 and pAcC12-GAP101-7.

18. Baculovirus vectors comprising the DNA sequence of Claim 1.

19. A method of producing a polypeptide having GAP activity wherein said polypeptide has the amino acid sequence shown in FIG.5, comprising the steps of:

(a) transforming host cells with said baculovirus transfer vector of Claim 16, or said baculovirus of Claim 18;

(b) amplifying said transformed host cells; and

(c) purifying GAP from said host cells.

20. GAP produced by the method of Claim 19.

21. A recombinant DNA molecule which encodes a GAP peptide and has the sequence:

5' AAAACTCATG CAAGGGAAGG GCAAAACCCA GTATGGTCAG
AAGAGTTTGT CTTTGATGAT CTTCCTCCTG ACATCAATAG
ATTTGAAATA ACTCTTAGTA ATAAAACAAA GAAAAGCAAA
GATCCTGATA TCTTATTTAT GCGCTGCCAG TTGAGCCGAT
TACAGAAAGG GCATGCCACA GATGAATGGT TTCTGCTCAG
CTCCCATATA CCATTAAAAG GTATTGAACC AGGGTCCCTG
CGTGTTCGAG CACGATACTC TATGGAAAAA ATCATGCCAG
AAGAAGAGTA CAGTGAATTT AAAGAGCTTA TACTGCAAAA
GGAACTTCAT GTAGTCTATG CTTTATCACA T 3'

22. The DNA molecule of Claim 20 wherein said DNA is of human origin.

23. Host cells transformed with the DNA molecule of Claim 21.

24. GAP 2.

25. GAP 6.

26. GAP 8.

27. A method of diagnosing for cancer comprising detecting GAP gene amplification in cells suspected of being cancerous, comprising the steps of:

(a) immobilizing DNA from said cells on a solid support;

(b) subjecting said immobilized DNA to hybridization with a labelled GAP DNA probe; and

(c) determining the amount of labelled GAP probe bound to said cell DNA.

28. The method as described in Claim 26 wherein said labelled GAP DNA probe is selected from the group consisting of GAP 2, GAP 6, GAP 8, clone 101, clone 7, clone 16 and clone "Sleepy", or fragments having a hybridizable sequence derived therefrom.

29. A method of diagnosing for cancerous cells, comprising the steps of:

(a) isolating mRNA from said cells suspected of being cancerous, said mRNA comprising mRNA complementary to labelled GAP DNA sequences;

(b) hybridizing said labelled GAP DNA sequences to said mRNA complementary to said GAP sequences; and

(c) determining the level of binding of said labelled GAP DNA.

30. The method as described in Claim 29 wherein said mRNA is poly (A+) mRNA.

31. The method as described in Claim 29 wherein said labelled GAP DNA sequences are selected from the group consisting of GAP 2, GAP 6, GAP 8, clone 101, clone 7, clone 16, and "Sleepy", or hybridizable fragments derived therefrom.

32. A recombinant DNA molecule encoding a polypeptide possessing GAP activity comprising a nucleotide sequence shown in FIG. 5, wherein the nucleotides that encode the first 180 amino acids are deleted.

33. A substantially pure protein molecule comprising an apparent molecular weight of about 115,000-120,000 daltons as assessed by sodium dodecyl sulfate polyacrylamide electrophoresis run under reducing conditions, or fragments derived therefrom, that stimulate non-oncogenic ras p21 GTPase activity.

34. A molecule as described in claim 33 wherein said fragments have an apparent molecular weight of about 45,000-55,000 daltons as assessed by sodium dodecyl sulfate polyacrylamide electrophoresis run under reducing conditions.

35. A molecule as described claim 34 wherein said fragments have an apparent molecular weight of about 50,000 daltons as assessed by sodium dodecyl sulfate polyacrylamide electrophoresis run under reducing conditions.

36. A method of purifying a molecule having an apparent reduced molecular weight of about 115,000-120,000 daltons that stimulates non-oncogenic ras p21 GTPase activity from a solution containing the same, comprising the steps of:

contacting said solution with cation exchange chromatographic material for a time sufficient for said molecule to bind to said material;

forming a first eluate containing said molecule by eluting said molecule from said cation chromatographic material by contacting said chromatographic material with an aqueous salt solution;

identifying fractions in said first eluate having said molecule, and reducing the salt concentration present in said fractions to be compatible with anion exchange chromatography;

forming a second eluate by contacting said fractions of said first eluate containing said molecule with anion exchange chromatographic material for a time sufficient for said molecule to bind to said material, and eluting said molecule from said material anion exchange material with an aqueous salt solution;

forming a third eluate by contacting said second eluate with a second cation exchange chromatographic material for a time sufficient for said molecule to bind to said material, and eluting said material from said second cation exchange chromatographic material; and

identifying fractions of said third eluate containing said molecule.

37. The method as described in claim 36 wherein said molecule is purified from human placenta.

38. The method as described in claim 36 wherein said purification is conducted in solutions containing one or more protease inhibitors at concentrations that effectively inhibit proteolysis of said molecule.

39. The method as described in claim 38 wherein said molecule is present in a solution comprising a reducing agent in an amount effective to preserve the activity of said molecule.

40. The method as described in claim 39 wherein said purification is conducted in solutions containing one or more metal ion chelators at concentrations that prevent substantial loss of activity of said molecule.

41. The method as described in claim 40 wherein said purification is conducted in solutions containing cations in an amount that substantially maintains the activity of said molecule.

42. The method as described in claim 41 wherein said cation is magnesium chloride.

43. The method as described in claim 36 wherein said solution contains about 1 mM magnesium chloride, 5 mM EGTA, 0.1 mM DTT, and 100 μM PMSF.

44. The method as described in claim 36 wherein said first cation exchange chromatographic material is S-Sepharose.

45. The method as described in claim 36 wherein said anion exchange chromatographic material is diethyl aminoethyl.

46. The method as described in claim 36 wherein said second cation chromatographic material comprises sulfopropyl functional groups.

47. The method as described in claim 46 wherein said second cation exchange material is TSK-SE-5-PW.

48. The method as described in claim 36 wherein a hydrophobic interaction chromatographic material is substituted for said second cation exchange material, and said second eluate from said anion exchange chromatographic material is made compatible with said hydrophobic interaction chromatographic material by increasing the ionic strength of said second eluate, and forming a third eluate by eluting said molecule from said hydrophobic interaction chromatographic material.

49. The method as described in claim 48 wherein said molecule is eluted from said hydrophobic interaction chromatographic material with an aqueous low salt solution, or an effective amount of a polyol.

50. The method as described in claim 48 wherein said hydrophobic interaction chromatographic material is under high pressure.

51. A method for identifying cancer therapeutics, comprising the steps of:

(a) combining in a GAP assay compatible solution a compound suspected of being a cancer therapeutic, ras p21 protein having GAP stimulatable GTPase activity, and labeled GTP;

(b) measuring the amount of GTP converted to GDP plus phosphate; and

(c) relating the amount of GTP converted to GDP with a control sample prepared in accordance with step "a", said control sample being known to be free of said suspected cancer therapeutic.

52. The method as described in claim 51 wherein said GTP is labeled.

53. The method as described in claim 51 wherein GAP has a reduced molecular weight of about 115,000-120,000 daltons.

54. The method as described in claim 51 wherein GAP comprises a molecule having a reduced molecular weight of about 45-55,000 daltons and is a fragment derived from a molecule having a reduced molecular weight of about 115,000-120,000 daltons.

55. Cancer therapeutics identified by the method of claim 51.

FIGURE 1

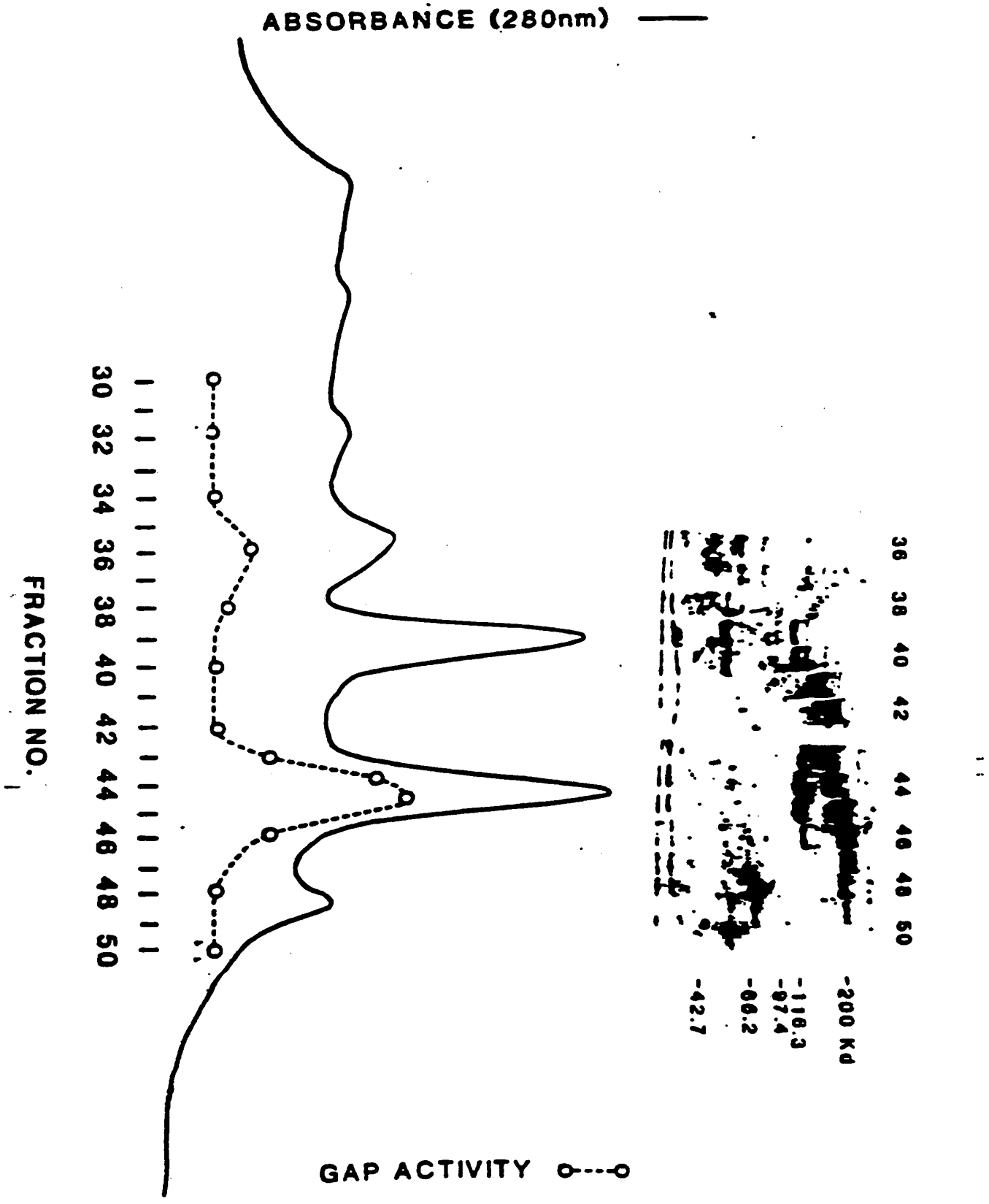
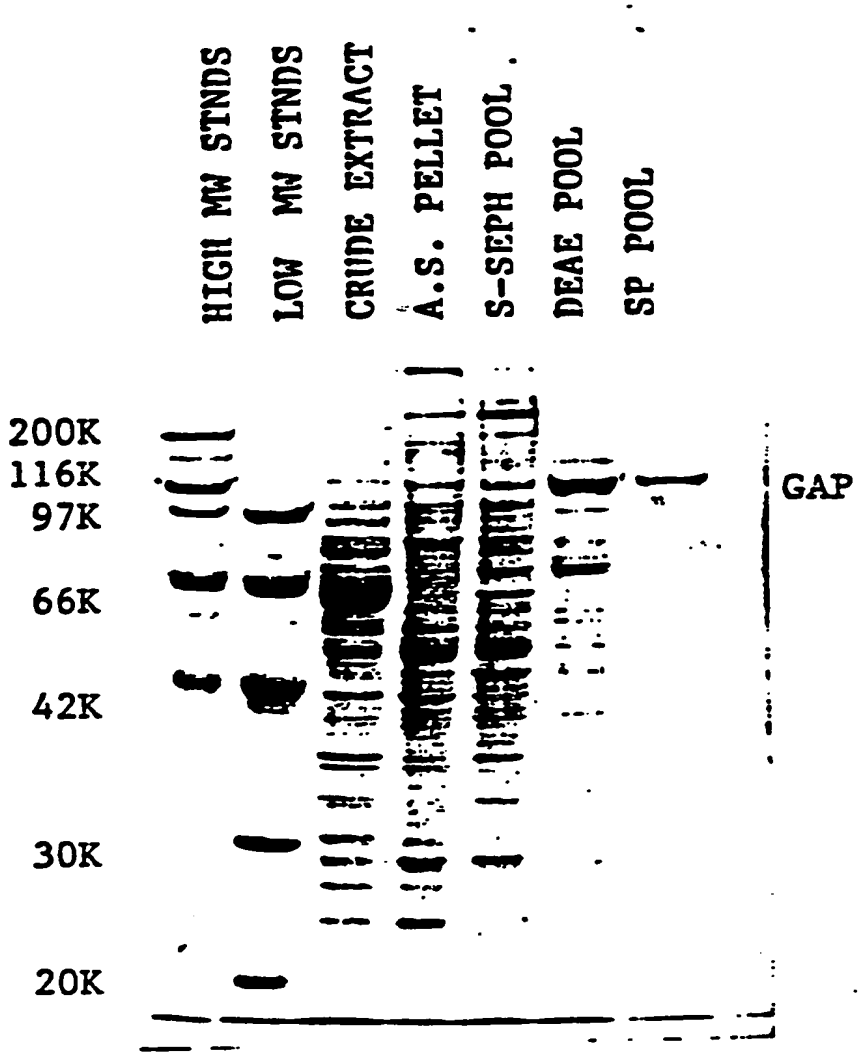


FIGURE 2



REDUCED 10% SDS PAGE
(COOMASSIE-STAINED)

FIGURE 3

ILE MET PRO GLU GLU GLU TYR SER GLU PHE LYS

```
    ATC ATG CCC GAG CAG GAG TAC TCC GAG TTC AAG
      T       A   AGA   A   T   T   A   T   A
5     A       T                   A
              G                   G
                                 AGC
                                   T
```

FIGURE 4

Ile MET Pro Glu Glu Glu Tyr Ser Glu Phe Lys

GW13 5'ATC ATG CCT GAG CAG GAG TAC TCT GAG TTC AAG'3

GW15 5'ATC ATG CCT GAG CAG GAG TAC AGT GAG TTC AAG'3

GW17 5'ATC ATG CCT GAG GAG GAG TAC TCT GAG TTC AAG'3

GW19 5'ATC ATG CCT GAG GAG GAG TAC AGT GAG TTC AAG'3

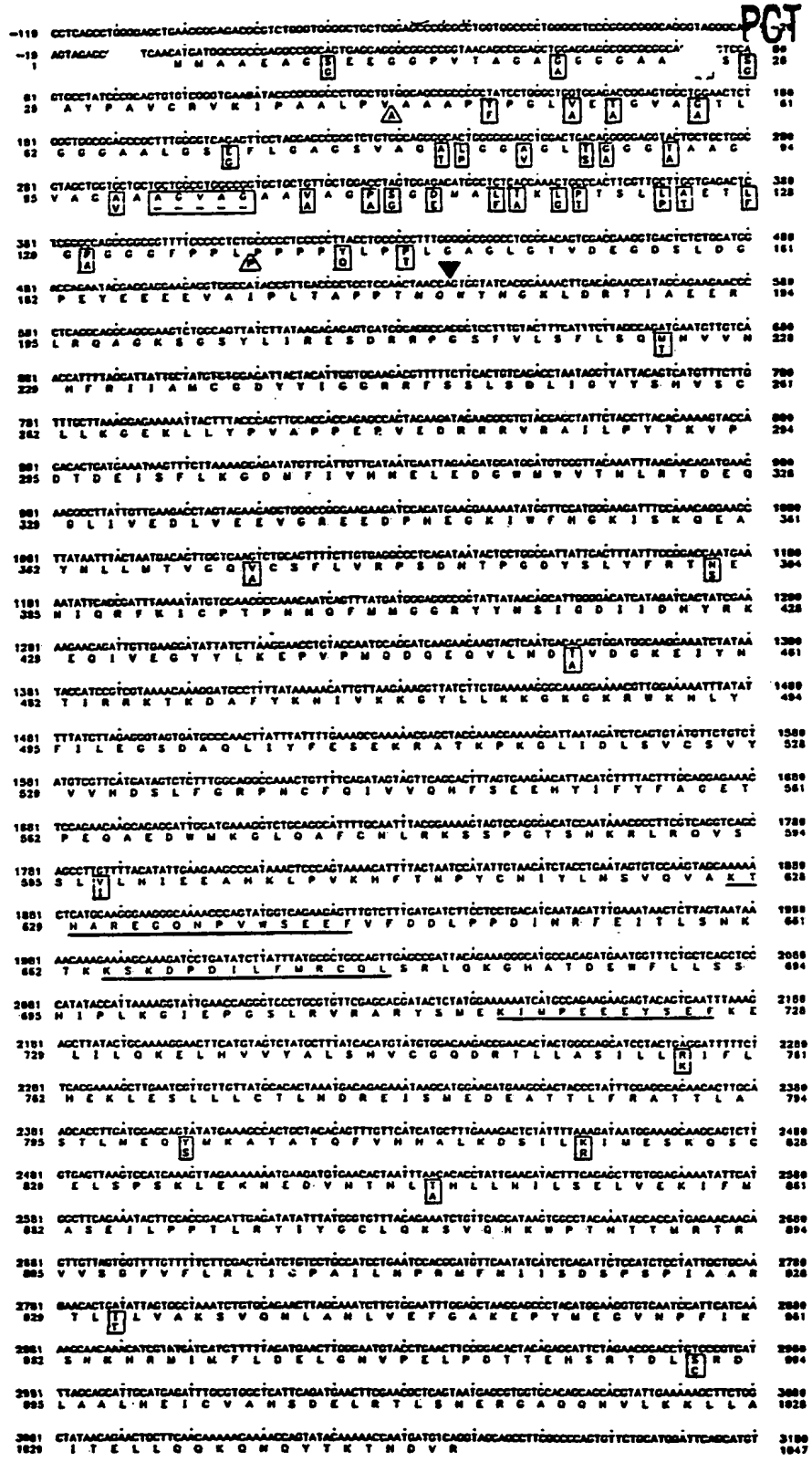


Fig C

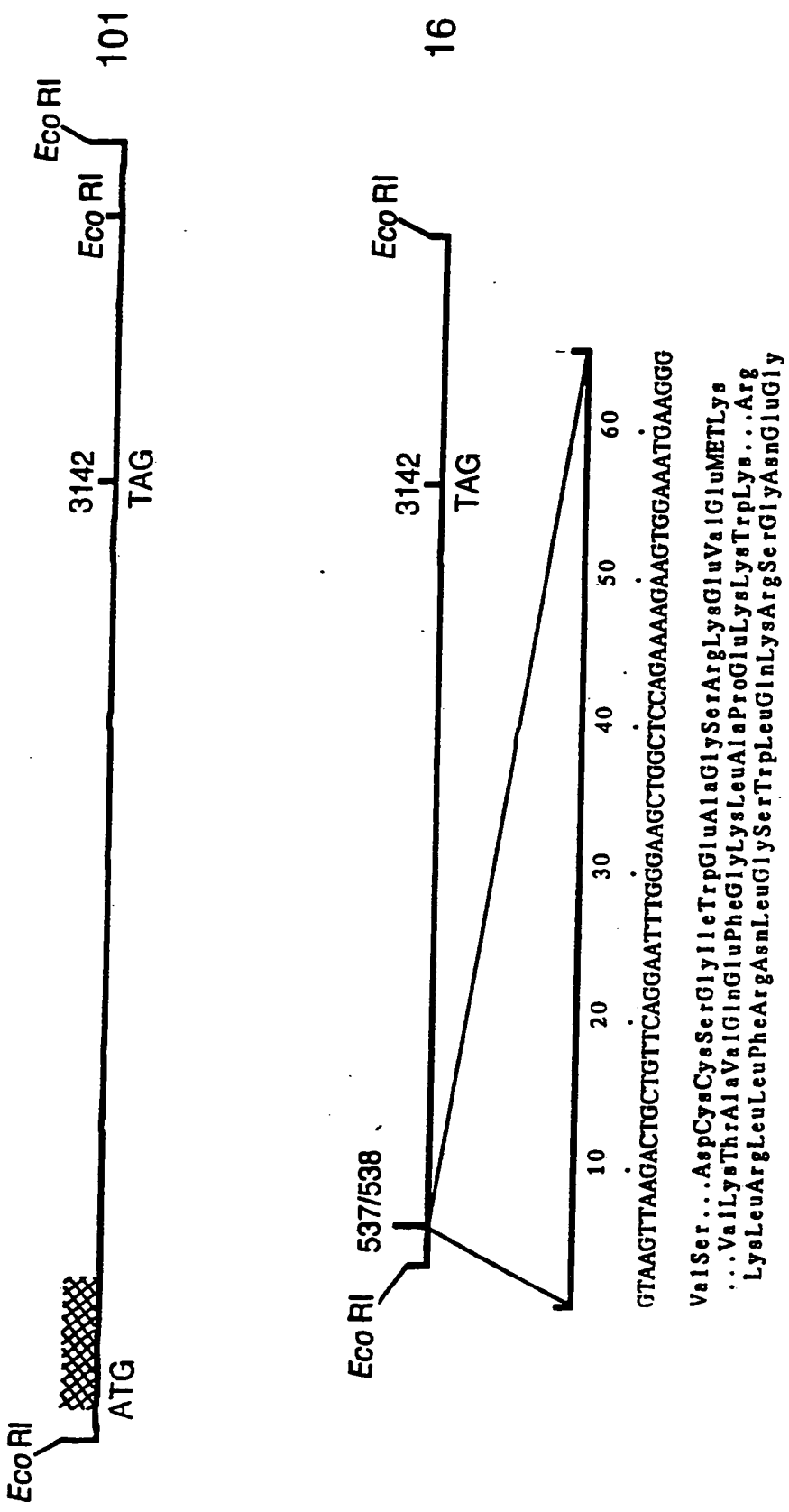


Fig. 6

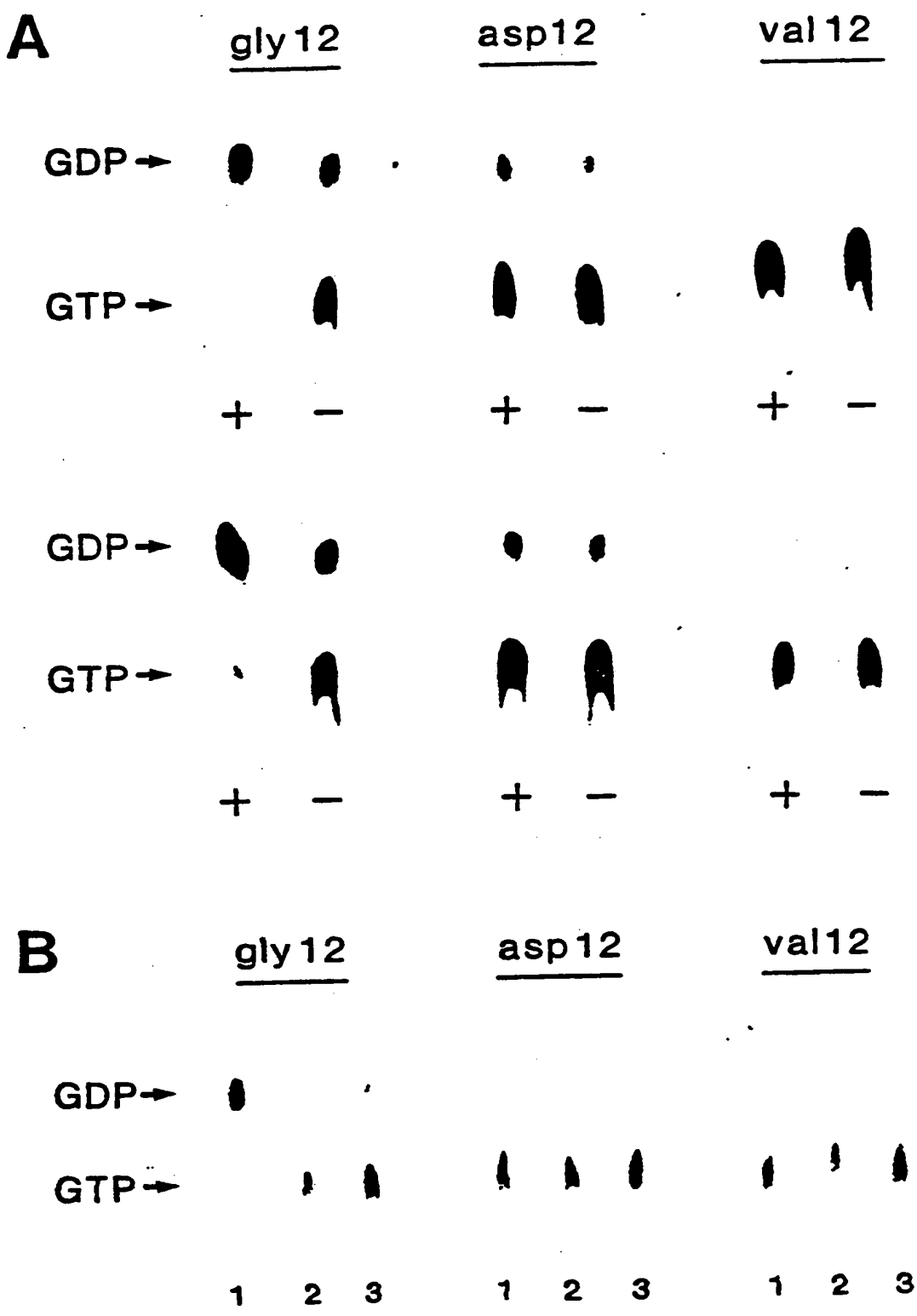
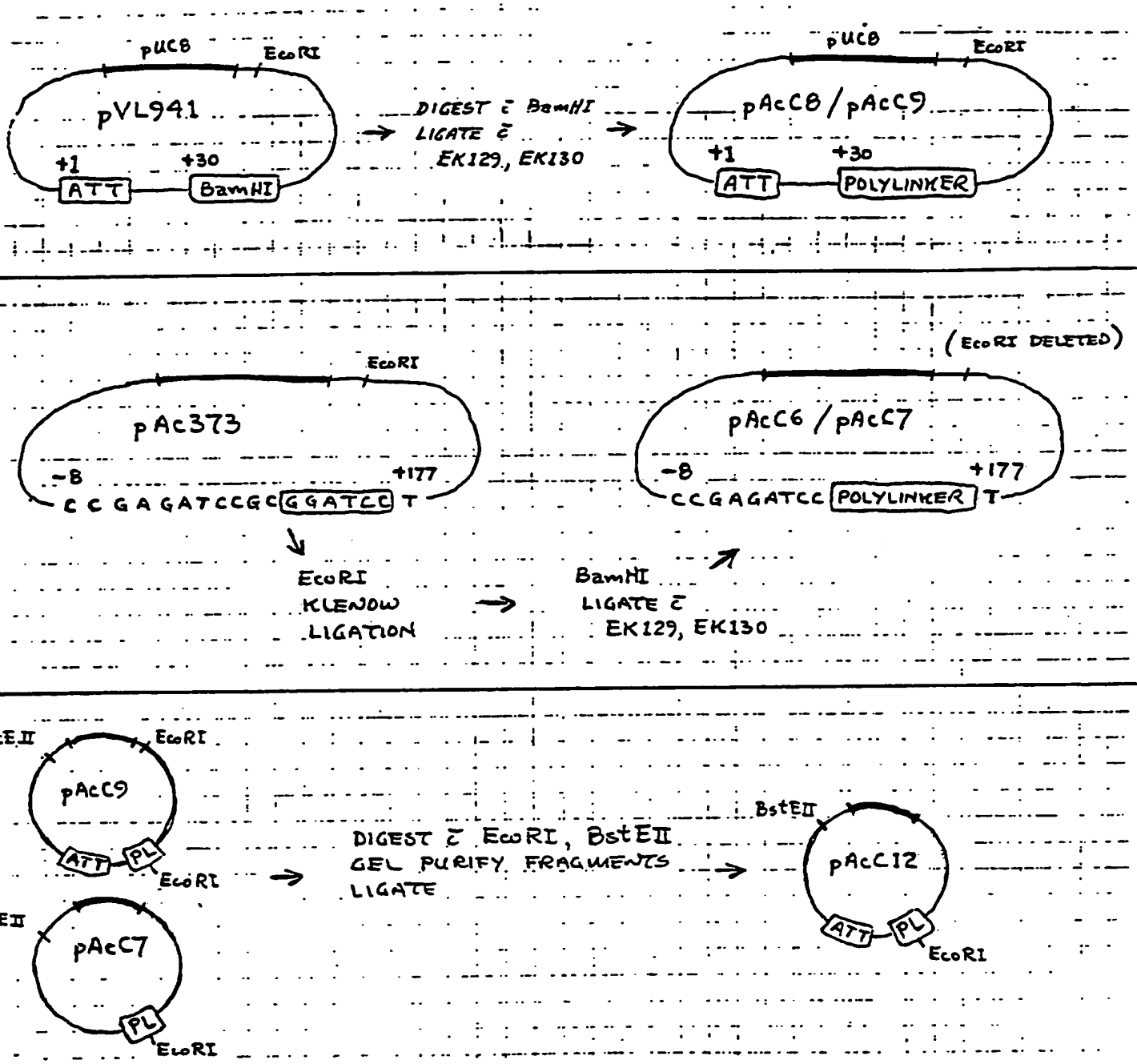


Fig. 7



ORIENTATION OF POLYLINKER

IN pAcC6, pAcC8 :

BamHI SacI BglII EcoRI SmaI
GATCCACCATGGAGCTCGAGATCTAGAATTCTGCAGCCCGGGGTACCGATC
NcoI XhoI XbaI PstI KpnI

IN pAcC7, pAcC9, pAcC12 :

SmaI EcoRI BglII SacI BamHI
GATCGGTACCCGGGCTGCAGAATTCTAGATCTCGAGCTCCATGGTGGATC
KpnI PstI XbaI XhoI NcoI

Fig 8

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/12, 15/86, C12P 21/02
C07K 13/00, C12Q 1/68, 1/42

A3

(11) International Publication Number: WO 90/00607

(43) International Publication Date: 25 January 1990 (25.01.90)

(21) International Application Number: PCT/US89/02761

(22) International Filing Date: 23 June 1989 (23.06.89)

(30) Priority data:

216,888	8 July 1988 (08.07.88)	US
230,761	10 August 1988 (10.08.88)	US
260,807	21 October 1988 (21.10.88)	US

(71) Applicant: CETUS CORPORATION [US/US]; 1400 Fifty-Third Street, Emeryville, CA 94608 (US).

(72) Inventors: WONG, Gail, L. ; 11255 Elvessa Street, Oakland, CA 94605 (US). MCCORMICK, Francis, P. ; 921 Ramona Avenue, Albany, CA 94706 (US). MARTIN, George ; 2346 Woolsey Street, Berkeley, CA 94705 (US). RUBINFELD, Bonnee ; 3403 Claridge Drive, Danville, CA 94526 (US). O'ROURKE, Edward, C. ; 7321 Skyline Boulevard, Oakland, CA 94611 (US). CLARK, Robin ; 3736 Woodruff Avenue, Oakland, CA 94602 (US). KOTHS, Kirston, E. ; 2646 Mira Vista Drive, El Cerrito, CA 94530 (US). HALENBECK, Robert, F. ; 136 Spring Grove Avenue, San Rafael, CA 94901 (US). TRAHEY, Mary, M. ; 5933 Chabot Road, Oakland, CA 94618 (US).

(74) Agent: HALLUIN, Albert, P.; Cetus Corporation, 1400 Fifty-Third Street, Emeryville, CA 94608 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), BG, CH (European patent), DE (European patent), DK, FI, FR (European patent), GB (European patent), HU, IT (European patent), JP, LU (European patent), NL (European patent), NO, SE (European patent).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
1 November 1990 (01.11.90)

(54) Title: GAP GENE SEQUENCES AND DIAGNOSTIC USES THEREOF

ILE MET PRO GLU GLU GLU TYR SER GLU PHE LYS

```
     ATC ATG CCC GAG CAG GAG TAC TCC GAG TTC AAG
       T       A   AGA   A   T   T   A   T   A
5      A       T                   A
               G                   G
                                  AGC
                                   T
```

(57) Abstract

Guanosine triphosphatase activating protein (GAP) DNA sequences are described that are useful as cancer diagnostics, particularly to detect cancer cells that express the ras oncogene protein p21 by measuring the level of GAP gene expression or amplification.

DESIGNATIONS OF "DE"

Until further notice, any designation of "DE" in any international application whose international filing date is prior to October 3, 1990, shall have effect in the territory of the Federal Republic of Germany with the exception of the territory of the former German Democratic Republic.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CM	Cameroon
DE	Germany, Federal Republic of
DK	Denmark
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GR	Greece
HU	Hungary
IT	Italy
JP	Japan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
MC	Monaco
MG	Madagascar
ML	Mali
MR	Mauritania
MW	Malawi
NL	Netherlands
NO	Norway
RO	Romania
SD	Sudan
SE	Sweden
SN	Senegal
SU	Soviet Union
TD	Chad
TG	Togo
US	United States of America

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 89/02761

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[5]: C 12 N 15/12, 15/86, C 12 P 21/02, C 07 K 13/00, C 12 Q 1/68, 1/42

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[5]	C 12 N, C 12 P, C 12 Q

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [10]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
A	Science, volume 238, 23 October 1987, (Washington, DC, US), M. Trahey et al.: "A cytoplasmic protein stimulates normal N-ras p21 GRPase, but does not affect oncogenic mutants", pages 542-545 --	
A	Science, volume 240, 22 April 1988, (Washington, DC, US), H. Adari et al.: "Guanosine triphosphatase activating protein (GAP) interacts with the p21 ras effector binding domain", pages 518-521 --	
A	Nature, volume 332, 7 April 1988, (London, GB), C. Calés et al.: "The cytoplasmic protein GAP is implicated as the target for regulation by the ras gene product", pages 548-551 --	

./.

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
24th August 1990	04. 10. 90
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	R.J. Eernisse

Form PCT/ISA/210 (second sheet) (January 1985)

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, ¹¹ with indication, where appropriate, of the relevant passages	Relevant to Claim No.
P,X	Science, volume 242, 23 December 1988, (Washington, DC, US), M. Trahey et al.: "Molecular cloning of two types of GaP complementary DNA from human placenta", pages 1697-1700 --	1-50
P,X	Nature, volume 335, 1 September 1988, (London, GB), U.S. Vogel et al.: "Cloning of bovine GAP and its interaction with oncogenic ras p21", pages 90-93 --	33,36
P,X	Biochemical and Biophysical Research Communications, volume 159, no. 3, 31 March 1989, Academic Press, Inc., (Duluth, Mn, US), T. Ueda et al.: "GTPase activating proteins from the smg-21 GTP-binding protein having the same effector domain as the ras proteins in human platelets", pages 1411-1419 --	33,36
P,X	Cancer Research, volume 49, 15 March 1989, (Philadelphia, PA, US), C. Duanmu et al.: "Tubilin-dependent hydrolysis of guanosine triphosphate as a screening test to identify new antitubulin compounds with potential as antimitotic agents: application to carbamates of aromatic amines", pages 1344-1348	51-55

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This International search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers ....., because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim numbers ....., because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claim numbers....., because they are dependent claims and are not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. ☒ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this international application as follows:

1. Claims 1-50
2. Claims 51-55

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

☐ The additional search fees were accompanied by applicant's protest.

☒ No protest accompanied the payment of additional search fees.